NOTE TO USERS

This reproduction is the best copy available.

UMI®

University of Alberta

Characterizing the Protein and DNA Interactions of the F plasmid DNA Binding Protein, TraM

by

Richard A. Fekete ©

A thesis submitted to the Faculty of Graduate Studies and
Research in partial fulfillment of the requirements for the
degree of Doctor of Philosophy

in

Molecular Biology and Genetics

Department of Biological Sciences

Edmonton, Alberta

Spring, 2001

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-60291-5

Canada

University of Alberta

Library Release Form

Name of Author: Richard Alfred Fekete

Tittle of Thesis: Characterizing the Protein and DNA Interactions
of the F plasmid DNA Binding Protein, TraM

Degree: Doctor of Philosophy

Year this Degree Granted: 2001

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

8A, 10050-118st.
Edmonton, Alberta
T5K 2M8

Thursday, March 22, 2001

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Characterizing the Protein and DNA Interactions of the F Plasmid DNA Binding Protein, TraM** submitted by **Richard Alfred Fekete** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy,** in Molecular Biology and Genetics.

Laura Frost
Laura S. Frost (supervisor)

Brenda Leskiw
Brenda K. Leskiw

Tracy Raivio
Tracy Raivio

Mark Glover
Mark Glover

Beth Traxler
Beth Traxler

Michael A. Pickard (Committee Chair)

Date March 12, 2001

Abstract

Many host- and plasmid-encoded proteins are required to efficiently transfer the *Escherichia coli* F plasmid. The latter are divided into five functional groups: pilus synthesis and assembly, surface exclusion, mating aggregate stabilization, gene regulation, and DNA metabolism. Proteins of the last group are TraI, TraY, TraD, and TraM. TraI is a relaxase and helicase, and requires IHF and TraY, which bind to *oriT*, for cleavage. TraD is an inner membrane protein and is thought to link *oriT*-bound proteins to the transfer apparatus in the membrane. TraM binds to three sites in *oriT*, and is required for transfer, but not for cleavage *in vitro*. Two of the TraM binding sites, *sbmA* and *sbmB*, autoregulate *traM* transcription, and the last site, *sbmC*, is more important for transfer.

Nicking assays performed on an F *traM* mutant show that TraM is not required for relaxase activity *in vivo*. Nicking and mobilization assays using chimeric plasmids constructed from F and the F-like plasmid R100-1, using F or R100-1 Tra proteins provided in *trans*, suggest that protein:DNA interactions at *oriT* provide much of the previously characterized plasmid specificity. Biochemical characterization of TraM indicate that TraM exists mainly as tetramers in solution, but dimers and tetramers bind to TraM binding sites. Experiments involving the yeast two-hybrid system identified two domains which participate in TraM multimerization, a central and a carboxyl-terminal domain. Deleting various regions of TraM in the TraM:GAL4 fusions suggests that each domain interacts with a similar domain in another molecule to form TraM multimers.

EMSA experiments determined that TraM binds to its binding sites cooperatively and these complexes are very stable once bound. Hydroxyl radical footprinting defined which bases are protected by the bound protein.

Using this data, the factors responsible for relaxation and transfer mutant phenotypes are summarized. The identity of the domain responsible for tetramerization of TraM is also proposed based on similarities to the well-characterized LacR protein. Finally, a model describing the mechanism by which TraM binds to *oriT* is formulated based on data from EMSA, sizing of DNA bound proteins, and hydroxyl radical footprinting.

Acknowledgements

Throughout my time at the University of Alberta I have had the chance to meet and befriend may people. Without these people I would not have made it through the many years in science.

Foremost, I would like to thank my supervisor, Laura Frost. For some reason she took interest in me in the beginning of my studies and supported me even before I came to her lab for graduate studies. I would like to thank her for believing in me and supporting me at all times. During my research she let me pursue all my ideas, many of which came to nothing, however she still let me continue, hopefully knowing that something would eventually work. Her encouragement allowed me to venture to places I was unsure of and allowed me to meet many people in our field of work, and is something that has shaped me into becoming a scientist.

My committee has also greatly helped me throughout my studies, and so I owe much appreciation to Brenda Leskiw, George Owttrim, Mark Glover, and Linda Reha-Krantz. My thanks also go to Les Hicks in the Department of Biochemistry and Jack Moore at Alberta Peptide Institute for their help.

I would like to thank all of the members of our lab, past and present. Karen Anthony and Lori Jerome were good role models for me and gave me something to look up to. Jan Manchak kept us running all of the time and without her, many of my experiments would still be sitting in tubes. Mike Gubbins and the newer members of the lab have always provided us with comic relief and "interesting" discussions which took my mind off of work when I needed it the most. I am also thankful to Troy Locke for his help during protein purification. Most of all in our lab, I owe much gratitude to Bill Klimke. Starting at the same time as I did and coming from the same undergraduate program allowed us to think in the same way and gave me someone who I could relate to. Discussions about everything (with his humorous dismissive attitude) were things I really enjoyed with Bill. Outside the lab we also were friends giving me someone who was interested in the same things I was.

I would also like to thank my friends Jeff Cooper and Jay White who provided me with an outlet for many of my other feelings that normally built up. I have known Jeff since high school and he will always remain a friend. Jay I met at university and made a triplet with Bill and I (even though we don't look the same) that many people will not forget.

I would like to thank my parents for all of their support over the years. Even though they all knew that university was not easy, and helped me as much as possible, without my father, mother, and stepfather my university experience would not have been as enjoyable.

I would like to thank Christie Hamilton, who is more than a friend to me. She has listened to my ranting about everything and supported me every time I needed help. Her effort in getting my thesis together probably best showed how much I need her help. I owe her my gratitude and offer my love in return. Hopefully, I can give her back the help that she has given me.

Table of Contents

Chapter **Page**

1 **General Introduction** 1
- DNA in the environment 2
- DNA replication 3
  - Rolling circle-type replication 3
    - pT181 4
    - pC194 8
    - pMV158 8
    - ϕX174 9
  - Theta-type replication 10
    - DNA polymerase I dependent 10
    - DNA polymerase I independent 11
    - *oriC* 12
    - F plasmid 13
- Horizontal gene transfer 14
- Conjugation 15
- F Tra proteins 19
- Tra proteins involved in DNA metabolism 20
  - TraY 20
  - TraI 21
  - Relaxases 23
  - TraD 25
  - TraM 26
- Objectives 29

2 **Materials and Methods** 31
- Bacterial strains 32

Recombinant DNA techniques 32
Chimeric plasmid construction 33
Plasmid nicking assays 34
Nicking efficiency quantification 35
Mobilization efficiency assays 35
Cell fractionation 36
Immunoblot analysis 37
Construction of pRFM2T75 37
Construction of His-tagged TraM 38
Overexpression and purification of TraM 38
Column chromatography 39
Protein analysis 40
Size quantification of TraM 41
Construction of TraM binding site clones 42
Chemical crosslinking 43
Determination of the oligomeric state of TraM bound to DNA 44
Construction of plasmids for yeast two-hybrid system 47
Liquid media and plates for yeast two-hybrid system 49
Transformation of yeast 49
β-galactosidase assays 50
Confirmation of yeast two-hybrid system mutants 51
Electromobility shift assays 51
Hill plots 52
Bending and competition assays 53
Hydroxyl radical footprinting 54

3 **Analysis of the allele specific nature of *oriT*** 56
Introduction 57
Results 61
F plasmid nicking assay 61
Chimeric plasmids 64

Chimeric plasmid mobilization assays 64

Chimeric plasmid nicking assays 70

Protein localization of TraM 74

Discussion 82

4 Biochemical characterization of TraM in solution 88

Introduction 89

Results 95

Overexpression and purification 95

Mass spectrometry 95

Analytical ultracentrifugation 100

Size exclusion chromatography 103

Chemical crosslinking 106

Non-denaturing gel electrophoresis to size bound TraM 109

Discussion 119

5 Identification and analysis of the TraM multimerization domains using the yeast two-hybrid system 124

Introduction 125

Results 134

Wild-type TraM fusions 134

TraM:GAL4 activating domain deletions 135

TraM:GAL4 binding domain deletions 140

Discussion 151

6 Characterization of the DNA binding activity of TraM 155

Introduction 155

Results 163

EMSA of TraM binding sites 163

Characterization of single-stranded *sbmA* 178

Measurements of cooperativity 187
DNA bending by bound TraM 195
Competition assays 201
Hydroxyl radical footprinting 209
TraM binding site sequence alignments 219
Discussion 226

7 **General Discussion** 232
TraM specificity 223
TraI specificity 234
TraM multimerization 237
DNA binding of TraM 242
Base specific DNA contacts 244
Model for TraM binding 245

8 **Appendix** 250
Creating an epitope tagged TraM fusion 251
Effects of running conditions on EMSA 252
sbmAB circularization experiments 255
Binding of F TraM to F and R100-1 TraM binding sites 256

9 **References** 262

List of Tables

Table	Title	Page
3.1A	Mobilization of chimeric plasmids	68
3.1B	Effect of chimeric plamsids on transfer of pOX38-Km, pOX-*traM*K3	68
3.2	Percentage of chimeric plasmids cleaved at *nic*	73
3.3	Glucose-6-phosphate dehydrogenase activity of cell fractions	78
4.1	Size of protein complex bound to TraM binding sites	117
6.1	Sequences of the TraM binding sites cloned into pBEND2	165
6.2	Quantitative analysis of TraM binding to the TraM binding sites	177
6.3	Oligonucleotides used in EMSA of TraM	181

List of Figures

Figure	Title	Page
1.1	Replication cycle of rolling circle replicating plasmid pT181	7
1.2	Diagrams of the transfer regions of F, R100-1, and R1	18
1.3	Protein sequence comparison of TraM from F, R100-1, and R1	28
3.1	Nicking reactions of pOX38-Km and its derivatives	63
3.2	Sequences of the *oriT* regions of the chimeric plasmids	66
3.3	Examples of nicking reactions using the chimeric plasmids	72
3.4	Immunoblot to determine the presence of TraM in all chimeric plasmids	76
3.5	Protein localization by immunoblot analysis	80
4.1	Summary of steps involved in the purification of TraM	97
4.2	Mass spectrum of F TraM dissolved in TED	99
4.3	TraM distribution during analytical ultracentrifugation	102
4.4	Standard curve used during size exclusion chromatography	105
4.5	Immunoblots of TraM after crosslinking	108
4.6	DNA bound TraM run on a non-denaturing polyacrylamide gel	111
4.7	Graph of the relative mobility of protein standards used to create a standard curve	114
4.8	Graphs of the relative mobility of free and protein bound TraM binding sites	116
5.1	The reporter genes in the yeast strain PJ69-4A	129
5.2	Representation of the GAL4:TraM fusions	137
5.3	β-galactosidase activity of the activating domain constructs	139

5.4 Helical wheel analysis of TraM 142
5.5 Growth of yeast strains on drop-out medium 144
5.6 β-galactosidase activity of the binding domain constructs 147
5.7 β-galactosidase activity of all plasmid combinations 149

6.1 The bend angle effects end to end distance of DNA molecules 160
6.2 Binding of TraM to *sbmA*, *sbmC*, *sbmAB*, and *sbmABC* 167
6.3 Binding of TraM to *sbmAA*, half *sbmA*, and a negative control 169
6.4 Finer dilutions of TraM binding to *sbmA*, negative control, and *sbmC* 171
6.5 Finer dilutions of TraM binding to *sbmAB* and *sbmABC* 173
6.6 Graphs to determine association constants 176
6.7 The buildup and loss of the various *sbmAB* complexes 180
6.8 Binding of TraM to single-strands of the *sbmA* site 183
6.9 Characterization of TraM binding to a modified *sbmA* site 185
6.10 Analysis of the ability of TraM to bind *sbmA* half sites 189
6.11 Hill plots of the TraM binding sites 192
6.12 Bending of *sbmA* and *sbmC* upon TraM binding 197
6.13 Bending of *sbmAB* and *sbmABC* upon TraM binding 199
6.14 Competition assays of TraM bound to *sbmA* and *sbmC* 203
6.15 Competition assays of TraM bound to *sbmAB* 206
6.16 Competition assays of TraM bound to *sbmABC* 208
6.17 Hydroxyl radical footprinting of TraM bound to *sbmA* 211
6.18 Mapping the protected bases on *sbmA* 213
6.19 Hydroxyl radical footprinting of TraM bound to *sbmC* 216
6.20 Mapping the protected bases on *sbmC* 218
6.21 Consensus sequences for TraM binding sites 221
6.22 TraM consensus sequences in *oriT* 224

7.1 Conjugation steps mediated by TraM, TraY, TraI, and TraD 236

7.2 The four-helix bundle structure of the Lac repressor 241
7.3 Model showing how TraM binds to its binding sites 247

8.1 Effects of current on EMSA of *sbmAB* 254
8.2 Detection of circular *sbmAB* species using non-denaturing gels 258
8.3 EMSA of F TraM using F and R100-1 TraM binding sites 261

Abbreviations

aa	Amino acid
A	Adenine
Amp	Ampicillin
AP	Alkaline phosphatase
APS	Ammonium persulfate
ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine serum albumin
BPB	Bromophenol Blue
C	Cytosine
cpm	Counts per minute
Da	Daltons
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
ds	Double-stranded
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electromobility shift assays
F	Fertility sex factor
g	Gravity
g	Gram
G	Guanine
Glc	Glucose
hr	Hour
HRP	Horse radish peroxidase
IHF	Integration host factor
IgG	Immunoglobulin G
IPTG	Isopropyl β-D-thiogalactopyranoside
K_a	Equilibrium association constant
Kan	Kanamycin
kbp	Kilobase pairs
K_d	Equilibrium dissociation constant
kDa	Kilodalton
L	Litre

LB	Luria-Bertani medium
M	Molar
mCi	Millicurie
min	Minute
mA	Milliampere
mM	Millimolar
MU	Miller Units
NaI	Naladixic acid
nM	Nanomolar
O.D.$_{600}$	Optical Density at 600 nm
ONPG	2-Nitro-phenyl-β-D-galactopyranoside
ORF	Open reading frame
oriT	Origin of transfer
PCR	Polymerase chain reaction
pg	Picogram
pM	Picomolar
PVDF	Polyvinylidene difluoride
^{R}	Resistant
rpm	Revolutions per minute
RBS	Ribosome binding site
SDS	Sodium dodecylsulfate
SDS-PAGE	Sodium dodecylsulfate-polyacrylamide gel electrophoresis
Sp	Spectinomycin
ss	Single stranded
T	Thymine
Tc	Tetracycline
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
V	Volts
w/v	Weight/volume
µL	Microlitre
^{o}C	Degree Celsius

Chapter 1

Introduction

General Introduction

DNA in the environment

The ability of bacteria to exploit new environments has been attributed to horizontal transfer of DNA, rather than modification of existing genes (Syvanen, 1994). This gene transfer has been observed in a wide range of organisms, including transfer from bacteria to plants and yeast, and also from plants to bacteria (reviewed in Davison, 1999). Comparison of the *Escherichia coli* and *Salmonella enterica* genomes has suggested that approximately 17.6% of the genes present in the *E. coli* genome have been acquired by horizontal transfer. The ubiquitous nature of gene transfer between species, genera and kingdoms and evidence showing that transfer has occurred illustrates the importance of horizontal transfer. This is particularly true when looking at the deliberate and accidental release of recombinant and non-recombinant organisms into the environment. Often these organisms would not be able to compete with established organisms, but it is possible that they could release new genes into the environment.

Limitations to transfer are thought to control the rate of gene transfer. These limitations include proper environment for transfer, DNA restriction enzymes in the recipient, and stable maintenance of the DNA in the recipient. Restriction enzymes can be controlled using restriction protection systems that protect the transferred DNA (Chilley and Wilkins, 1995). Stable maintenance of the transferred DNA can be aided by its ability to confer a selective advantage to the host such as antibiotic resistance, heavy metal resistance or the ability to catabolize new compounds. However, the most important factor in the maintenance of transferred DNA is its ability to be replicated in the recipient. This can be accomplished by autonomous replication of the transferred DNA or by recombination into the recipient genome.

DNA replication

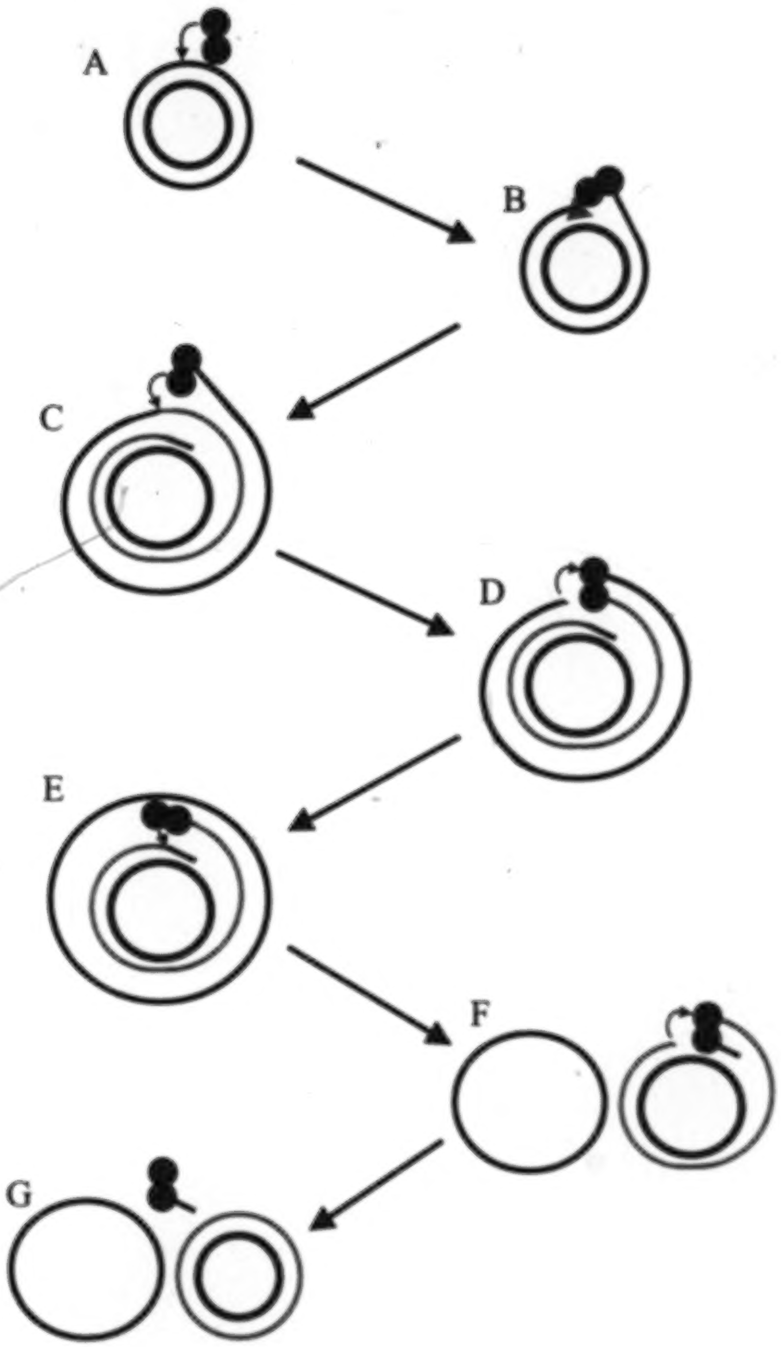
Many types of replication have been described, however, two common types are theta replication (including the *E. coli* chromosome) and rolling circle replication (RCR). Rolling circle replicating plasmids are found in Gram-positive and Gram-negative bacteria, and replication occurs unidirectionally with leading- and lagging-strand synthesis uncoupled (reviewed in del Solar *et al.*, 1993b). The replication proteins (Rep) cleave a double-stranded origin (*dso*), leaving a free 3' hydroxyl which is used as a substrate for leading-strand synthesis. DNA synthesis occurs until the *dso* is reached, producing a double-stranded plasmid and a displaced single-stranded intermediate (which is the hallmark of rolling circle replication; te Riele *et al.*, 1986). The single-stranded intermediate is then converted to double-stranded using a single-stranded origin (*sso*) of replication. The rate of conversion from single-stranded to double-stranded form (effecting the amount of single-stranded species present) depends on the efficiency with which the *sso* is recognized by host proteins (del Solar *et al.*, 1987). All polymerization steps for both strands use only host-encoded proteins. To initiate polymerization, only the plasmid-encoded leading-strand initiation and control (LIC) region is required. This region includes the *dso* and the *rep* gene (del Solar *et al.*, 1993a). Based on LIC sequence similarities RCR plasmids have been grouped into four families, represented by pT181 (Novick, 1989), pC194 (Gruss and Ehrlich, 1989), pMV158 (del Solar *et al.*, 1993b), and pSN2 (Novick, 1989). However, very little information is available on the last group. Similarity in the LIC of each family is very high, but drops outside this region. Two regions in the *dso*, *bind* and *nic* are used for Rep protein binding and cleavage, respectively (del Solar *et al.*, 1998). These two regions can be contiguous as in the pT181 family, or separated by up to 100 base pairs as in the pMV158 family. The *nic* sites are highly conserved within a family, whereas some variation is found within the *bind* sites, suggesting that the Rep proteins within a family share a common catalytic domain for cleavage, and plasmid specificity is provided by the *bind* loci. The *nic* sites contain an inverted repeat that is important for hairpin formation, which is required for cleavage (Noirot *et al.*, 1990). Termination requires only this hairpin (*nic* site) and does not need the *bind* loci (Zhao and Khan, 1996).

pT181 RCR plasmid family

The pT181-family *dso* is found within the *rep* coding region and contains two inverted repeats required for function: IR-II (containing the *nic* site) and IR-III (containing the *bind* site). The proximal arm of IR-III and spacing between the repeats are important for function (Wang *et al.*, 1993). Rep proteins from this family are more than 300 amino acids and RepC from pT181 is thought to function as a dimer (Jin *et al.*, 1996). The carboxyl-terminus of RepD from pC221 (in the pT181-family) interacts with IR-III (*bind*) and provides plasmid specificity (Thomas *et al.*, 1995). Six amino acids in the carboxyl-terminus of pT181 RepC are required for *bind* interaction (Wang *et al.*, 1992). The separation of the binding and nicking domains suggest that the protein is folded to bring these two domains together. Rep protein binding is also thought to cause unwinding at *nic* and/or cruciform formation in this family (Noirot *et al.*, 1990; Wang *et al.*, 1992). To initiate replication the RepC dimer binds to the *dso*. One subunit of the RepC dimer then cleaves a 5'-ApT-3' phosphodiester bond (Figure 1.1a) and tyrosine 191 of RepC in pT181 is covalently bound to the free 5' end (Thomas *et al.*, 1995) leaving the free 3' hydroxyl to be used as the substrate for replication (Figure 1.1b). After replication has proceeded for more than one round the second RepC subunit cleaves the first *nic* (where replication began) producing a free 3' hydroxyl group (Figure 1.1c). This performs a nucleophilic attack on the tyrosyl-phosphodiester bond of subunit one (Figure 1.1d), generating a single-stranded circular molecule. Subunit one then performs another nucleophilic attack on the newly synthesized *nic* (Figure 1.1e), covalently attaching to an oligonucleotide representing the 3' half of IR-II. This newly generated 3' hydroxyl then attacks the tyrosyl-phosphodiester bond of subunit 2 (Figure 1.1f) and forms a double-stranded covalently closed circular plasmid and a RepC dimer bound to an oligonucleotide (Figure 1.1g). This RepC heterodimer is unable to initiate replication again (Jin *et al.*, 1997) and ensures that a dimer is used only once per replication cycle.

After generation of the 3' hydroxyl group pT181 requires PcrA in *S. aureus* (the helicase II homologue in *E. coli*; Iordanescu, 1993), and it is thought that polymerase III extends the leading-strand. The SSB (single-strand binding protein) is also thought to be required to bind the displaced single-stranded intermediate. Lagging-strand synthesis is initiated at single-stranded origins (Novick, 1989). These origins are host-dependent and often require RNA polymerase and DNA polymerase I.

Figure 1.1. Replication cycle of rolling circle replicating plasmid pT181. The replication protein (RepC) dimer is shown as two red circles. The unnicked strand used as a template for replication is shown in blue with the nicked strand shown in black. The replication apparatus (and direction) is shown with an orange arrow, with the first DNA strand created from replication shown in orange, and the DNA created after a second pass of *nic* shown in pink. Nucleophilic attacks are shown with green arrows. A. Shows nucleophilic attack of *nic* by subunit 1 of RepC. B. Shows covalent attachment of the cleaved strand to subunit 1 of the RepC dimer. C. Shows the plasmid after more than one complete round of replication. D. After cleavage of the second *nic* by subunit 2 of RepC dimer, the newly freed 3' hydroxyl attacks tyrosyl-phosphodiester bond of RepC subunit 1, freeing a single-stranded plasmid (E). E. Newly freed RepC subunit 1 then attacks the *nic* site created from more than one full round of replication, covalently attaching itself to the oligonucleotide representing half of the *nic* site. F. The freed 3' hydroxyl attacks the tyrosyl-phosphodiester bond between subunit 2 of the RepC dimer creating a double-stranded covalently closed circular plasmid (G).


A
B
C
D
E
F
G

pC194 RCR plasmid family

The pC194-family has the *dso* within the Rep mRNA leader region. These *dso* are similar to the origins of the bacteriophage ϕX174, however, a stem loop is still formed at *nic* (Gruss and Ehrlich, 1989). Rep proteins from this family are approximately 300 amino acids and tyrosine 214 from RepA from pC194 is thought to cleave *nic* and covalently attach to the DNA (Noirot-Gros *et al.*, 1994). Glutamic acid 142 and 210 then contribute to termination by the hydrolysis of the newly synthesized *nic* using an activated water molecule, leaving the protein unmodified. Replacing glutamic acid 210 with tyrosine changed the reaction from hydrolysis to transesterification, allowing the reinitiation of replication in a ϕX174 manner (Noirot-Gros *et al.*, 1996).

pMV158 RCR plasmid family

The pMV158-family also has two inverted repeats, one at *nic* and the other at *bind*, separated by a spacer region of 14 to 95 base pairs. These plasmids contain the *dso* upstream of the *rep* gene, and encode Rep proteins of approximately 200 amino acids. Rep proteins in this family, such as RepB of pMV158, do not covalently bind to cleaved origins as do the Rep proteins of the pT181 and pC194-type plasmids (Moscoso *et al.*, 1995). However, a transient covalent bond exists between tyrosine 99 and the pMV158 origin (Moscoso *et al.*, 1997). Sequence analysis of the Rep proteins from this family suggest that the cleavage domain is closer to the amino-terminus, while the DNA binding domain is located in the carboxyl-terminus (del Solar *et al.*, 1993b).

ϕX174 RCR bacteriophage

The single-stranded bacteriophage ϕX174 has been very well characterized and uses a replication mechanism similar to that of RCR plasmids (reviewed in Baas and Jansz, 1988). Infection by a single-stranded DNA molecule is followed by conversion to double-stranded form using host proteins. During replication, the origin is cleaved and a mechanism similar to that of pT181 is used to spool off single-stranded DNA. The phage encoded gene A protein cleaves and covalently attaches to the 5' end of the DNA molecule leaving the 3' hydroxyl to be used as a substrate for DNA synthesis. Each protein molecule contains two tyrosines on the same side of an alpha-helix (van Mansfeld *et al.*, 1986) and either can perform the cleavage (Hanai and Wang, 1993). Only one protein molecule is covalently attached to the origin (Eisenberg and Kornberg, 1979) through a tyrosine-phophodiester bond (Roth *et al.*, 1984). After cleavage, gene A protein participates in the unwinding of the DNA (Ikeda *et al.*, 1976). After a full round of replication the second tyrosine performs a nucleophilic attack on the reformed origin (containing the portion used to prime DNA synthesis; van Mansfeld *et al.*, 1986), covalently linking itself to the DNA and producing a 3' hydroxyl end which attacks the tyrosyl-phosphodiester bond of the first tyrosine. This releases a circular single-stranded molecule that can be packaged. On a 30 nucleotide fragment containing the cleavage site between bases 7 and 8; bases 2 to 9 and 18 to 27 are bound by gene A protein (Heidekamp *et al.*, 1982). Between these sites is an A/T rich spacer which must be 8 bases; changing this spacing inhibits replication (Baas, 1987). This spacer also contributes to local unwinding when gene A protein is bound to the origin. Nucleotides 2 to 9 may only be recognized when nucleotides 18 to 27 are already bound, and possibly only when they are denatured (van Mansfeld *et al.*, 1980).

Theta-type plasmid replication: DNA polymerase I dependent

Theta-type replication is another common type of plasmid replication in bacteria. In this type of replication a bubble in the double-stranded DNA is formed where DNA replication occurs. This bubble grows larger as DNA is replicated. Theta replicating plasmids have been classified into two types: those that use DNA polymerase I and those that do not. DNA polymerase I dependent plasmids use a primer to initiate DNA polymerase I synthesis. For example, ColE1 origins contain two convergent promoters, from one of these RNAII is produced which primes DNA replication in a unidirectional manner. DNA polymerase I then performs a short synthesis which can be approximately 300 to 400 bases *in vitro* (Takechi *et al.*, 1995), but may be shorter *in vivo* (Janniere *et al.*, 1997). Replisome assembly involves loading the helicase and primase at primosome assembly sites found upstream of where DNA polymerase I is thought to arrest (Janniere *et al.*, 1997). This is followed by loading of DNA polymerase III, allowing leading and lagging-strand synthesis to occur.

Initiation of synthesis for all DNA polymerase I-dependent plasmids occurs in three steps: primer synthesis, DNA polymerase I entry and synthesis, and replisome assembly. The last two steps are similar for almost all DNA polymerase I-dependent plasmids. However, the first step has been used to place plasmids into five classes based on sequence and replication requirements (Espinosa *et al.*, 2000): ColE1, ColE2, pAMβ1, pCU1, and pJDB23 families. The ColE1 family includes the plasmids from which many cloning vectors have been derived (Baker and Wickner, 1992). ColE1 uses RNA polymerase and RNase H to produce a 3' hydroxyl group which can be used by DNA polymerase I (as stated above). Copy number control is accomplished using a convergent promoter to produce RNAI which prevents RNAII from forming the RNA/DNA hybrid required for DNA polymerase I. The Rom proteins (usually encoded downstream of the DNA polymerase I termination site) stabilize the RNA duplex, leading to its degradation. ColE2 plasmids use a Rep protein which binds to the origin and acts as a primase and is thought to provide the substrate needed for DNA polymerase I synthesis (Takechi *et al.*, 1995). The pAMβ1 family of plasmids has a Rep protein which binds to the origin, and a promoter upstream of the origin (possibly the *rep* promoter) which produces the substrate needed by DNA polymerase I, however, its formation is still unclear (Bruand *et al.*, 1993). Plasmid replication is controlled by the Cop protein, which represses the *rep* promoter, and an antisense RNA which causes premature termination of the *rep* transcript upstream of the open reading frame (Brantl, 1994; Brantl and Wagner, 1996).

Theta-type plasmid replication: DNA polymerase I independent

DNA polymerase I-independent theta-type replicating plasmids use a system similar to the replication of the *E. coli* chromosome from *oriC* (Kornberg and Baker, 1992). These plasmids have a replication initiator protein (Rep) which has binding sites at the origin (iterons). These origins also contain DnaA boxes, A/T rich regions (similar to the 13-mers at *oriC*), and GATC sites which are targets for the host Dam methylase. Examples of plasmids using this type of replication are F, P1, R1, and RK2/RP4. In general, the mechanism for replication begins with the binding of the Rep protein and DnaA, which causes a local melting of the A/T rich region. Interaction between DnaA and DnaC then allow loading of the helicase (DnaB), which further opens the replication bubble allowing loading of the primase (DnaG). Rep proteins also increase the ability of DnaB to load (Ratnakar *et al.*, 1996). Some plasmids, such as RSF1010, encode their own helicase (RepA) and primase (RepB) along with a replication initiator protein (RepC; Scherzinger *et al.*, 1991). After the primase is loaded, DNA polymerase III is loaded and proceeds in a continuous manner on the leading-strand and in a discontinuous manner on the lagging-strand. Replication can occur in a unidirectional or bidirectional manner from the origin. Using the Rep protein to initiate replication allows replication to be controlled by the transcriptional regulation of *rep* (Helinski *et al.*, 1996), the number of iterons (Park *et al.*, 1998), Rep protein dimerization (dimers can be nonfunctional; Ingmer *et al.*, 1995), and by handcuffing, which pairs plasmid molecules so that they cannot initiate (Pal and Chattoraj, 1988).

Escherichia coli chromosome: theta-type replication at *oriC*

Replication of the *E. coli* chromosome begins at *oriC* and progresses bidirectionally (DNA is replicated in both directions from the origin). The *oriC* contains four DnaA binding sites (9-mers which are called DnaA boxes) on one side of the origin and three 13-mers on the other side which can be bound by DnaA in a double-stranded or single-stranded form (Bramhill and Kornberg, 1988). The *oriC*, including these sites, is approximately 232 base pairs (Oka *et al.*, 1980), and insertions or deletions eliminate its function (Asada *et al.*, 1982). The 13-mers are located in an A/T rich region of approximately 45 base pairs, and once denatured are the site for DnaB (helicase) incorporation with the assistance of DnaC and DnaA. DnaA requires ATP to functionally bind and approximately 20-40 subunits form a complex with the DnaA boxes (Margulies and Kaguni, 1996). Once bound this complex causes strand opening in the A/T rich region. ADP blocks the ability of DnaA to cause strand opening, however, the ADP bound form is still able to bind *oriC* (Sekimizu *et al.*, 1987). Interaction of DnaA with the membrane speeds up the exchange of ADP for ATP, but can inhibit the ability of the protein to interact with *oriC* (Sekimizu *et al.*, 1988). DnaA binds to the DnaA boxes in a specific order, and is thought to occupy all of the sites only when replication is to be initiated (Margulies and Kaguni, 1996). Once all sites are occupied and the A/T rich region is denatured, DnaB is loaded with the assistance of DnaC and DnaA. DnaG (primase) and the DNA polymerase III holoenzyme are then loaded. The holoenzyme is made up of approximately ten different proteins: two core enzymes (containing the polymerase α or polC, the 5'-3' exonuclease ε or DnaQ, and θ), a τ or dnaX dimer (holding the two core enzymes together), a β or DnaN tetramer (responsible for clamping the complex to the DNA), and a γ complex composed of five different proteins (responsible for loading the β clamp; Marians, 1996). This complex allows leading and lagging-strand synthesis to occur simultaneously as the replication fork moves along the DNA in one direction.

F plasmid theta-type replication

The F plasmid contains three replication systems or replicons (Couturier *et al.*, 1988). The RepF1A replicon is of principal importance, since without it plasmids are much less stable. RepF1A has two origins, *oriV* and *oriS*, however *oriV* is non-essential. The Rep protein for *oriS* is RepE whose gene is found upstream of this origin. *oriS*, which induces theta-type replication, is approximately 217 base pairs in length and contains two DnaA boxes, a 46 base pair A/T rich region containing one 13-mer, and four 19 base pair RepE binding sites. Upstream of *oriS*, on the opposite side of *repE*, are five more RepE binding sites which define the incompatibility locus (*incC*) and are required for the proper regulation of initiation (Uga *et al.*, 1999, 2000). RepE is a dimer (Masson and Ray, 1988), however, replication also requires DnaA, B, C, G, E, RNA polymerase and DNA gyrase. It is thought that after opening of the A/T rich region one DnaB helicase is loaded and begins to unwind the plasmid on the discontinuous strand (Kornberg and Baker, 1992). A primosome assembly site (*pas*) is uncovered allowing continuous DNA replication to begin in the primary direction (Masai *et al.*, 1990). Another helicase is loaded at the *pas* and proceeds to unwind the plasmid in the opposite direction allowing continuous and discontinuous DNA replication to occur in both directions. This replication type has therefore been called sequentially bidirectional initiation (Kornberg and Baker, 1992).

Horizontal gene transfer

Horizontal transfer of genes has been divided into three groups based on the mechanism of transfer: transformation, transduction, and conjugation (Davison, 1999). Transformation is the uptake of naked DNA from the environment by a bacterium (Lorenz, 1994). DNA is present in the environment due to excretion or lysis of other cells. Transduction is the transfer of genes using a bacteriophage particle. Packaged genes can be generalized, where any genes are packaged (coliphage P1), or specialized, where DNA near the phage integration site in the genome is packaged (coliphage Lambda). Transduction is recipient-specific since the packaged phage must be able to infect the recipient cell for transfer to occur. Phage are prevalent in the environment, and phage encoding shiga toxin (thought to be involved in the pathogenicity of *E. coli* O157:H7) have been found in sewage at approximately 1-10 particles/mL (Muniesa and Jofre, 1998).

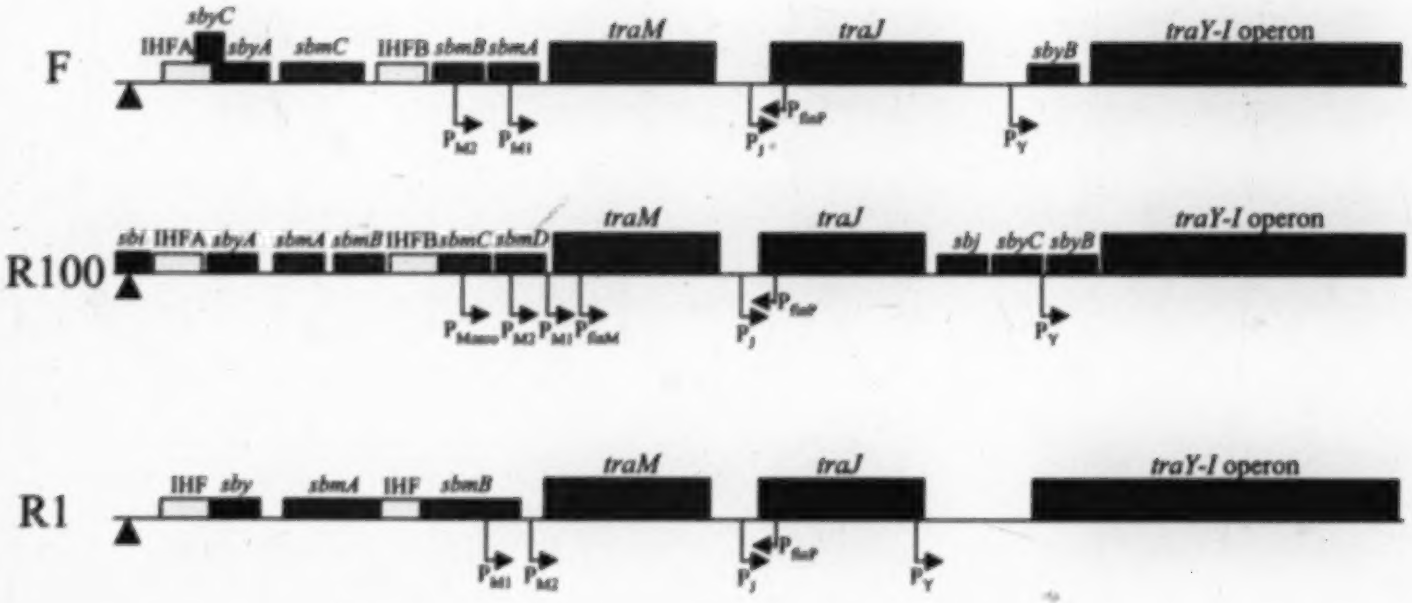
Conjugation

Conjugation is responsible for most horizontal gene transfer in the environment (Davison, 1999). It is the process by which DNA is transferred from a donor to a recipient cell through some type of mating pore in the donor and recipient cell membranes. The donor receives an unknown signal to initiate transfer, and in order to obtain efficient transfer the two cells must be stably attached. To mediate this attachment the donor cell synthesizes an extracellular filament called a pilus, which recognizes a potential recipient cell. After pilus retraction and cell:cell stabilization, a single-strand of DNA is transferred in a 5' to 3' direction through the mating pore into the recipient cell. DNA transfer starts from a site called *nic* in the origin of transfer (*oriT*). This is followed by complementary strand synthesis in the donor and recipient cells, after which both cells become transfer-competent donors. Conjugation can occur in a variety of ways: transfer of a self-transmissible plasmid like F or RP4, mobilization of plasmids like RSF1010 which contain a *nic* but do not encode a complete set of transfer genes, cointegration of a non-mobilizable plasmid into one that can be transferred, and transfer of chromosomal

DNA from an integrated self-transmissible plasmid. Conjugation has been shown to occur between a great variety of donors and recipients (reviewed in Davison, 1999) and has been suggested to be responsible for many clinical epidemics (Balis *et al.*, 1996). Transfer of genes involved in xenobiotic degradation (van der Meer, *et al.*, 1992), heavy metal resistance (Top *et al.*, 1990), and antibiotic resistance (Salyers and Shoemaker, 1996) are common and have been shown to occur in a variety of environments (Kruse and Sorum, 1994).

The *E. coli* F plasmid is the paradigm of self-transmissible plasmids. This 100 kbp circular plasmid belongs to the IncFI incompatibility group and is transferred in approximately 5 minutes. All plasmid-encoded proteins required for transfer are found in the 33.3 kbp transfer region (Figure 1.2). The organization of this region is similar to other IncF F-like plasmids such as R1 and R100 (Figure 1.2; Frost *et al.*, 1994) with the *nic* site at one end. Cleavage occurs on only one strand at *nic* (shown in Figure 1.2 as the lower strand). Based on the sequence similarity at *nic*, five families have been defined (Lanka and Wilkins, 1995; Guzman and Espinosa, 1997; Zechner *et al.*, 2000). Examples of each are RP4, F, RSF1010, ColE1, and pMV158 *nic* sites. Interestingly, comparing the origins used for DNA replication of the RCR plasmid pC194-family, the single-stranded bacteriophage ϕX174, and the transfer origins of IncP-like plasmids (RP4) gives a strong consensus of 5'-YAWCYTG*-3', where W represents A or T, Y represents pyrimidines, and * represents the cleavage site (Waters and Guiney, 1993; Pansegrau and Lanka, 1996a). To one side of the *nic* are the protein binding sites required for nicking and transfer of the plasmid. Together, these binding sites and *nic* are

Figure 1.2. Diagrams of the transfer regions of F, R100-1, and R1. Large boxes represent genes and smaller boxes represent the binding sites for those gene products. TraY binding sites: green; TraM binding sites: red; IHF binding sites: pink; TraJ binding sites: blue; TraI binding sites: black. *nic* sites are shown with large triangles below each figure, and the *tra* promoters are also designated with arrows. Not all regions are shown to scale.


F
IHFA
sbyC
sbyA
sbmC
IHFB
sbmB
sbmA
traM
traJ
sbyB
traY-I operon
P_{M2}
P_{M1}
P_{J}
P_{finP}
P_{Y}
R100
sbi
IHFA
sbyA
sbmA
sbmB
IHFB
sbmC
sbmD
traM
traJ
sbj
sbyC
sbyB
traY-I operon
P_{Mauto}
P_{M2}
P_{M1}
P_{finM}
P_{J}
P_{finP}
P_{Y}
R1
IHF
sby
sbmA
IHF
sbmB
traM
traJ
traY-I operon
P_{M1}
P_{M2}
P_{J}
P_{finP}
P_{Y}

called *oriT* (Willetts, 1972). This region is generally less than 500 bp in all transferable plasmids and removal of these sites causes significant decrease or elimination of transfer (Fu *et al.*, 1991). Further downstream in the F plasmid is the *traM* gene, followed by the *traJ* gene, and then the main *tra* operon which encodes the remainder of the plasmid-encoded genes required for conjugation. Since transfer occurs in a 5' to 3' direction this region is the last to be transferred during conjugation.

F Tra proteins

F plasmid-encoded proteins required for conjugation have been divided into 5 categories based on function: pilus synthesis and assembly, surface exclusion, mating aggregate stabilization, gene regulation, and DNA metabolism (Frost *et al.*, 1994). TraL, E (Frost *et al.*, 1984), K (Penfold *et al.*, 1994), C (Schandel *et al.*, 1992), and G (Firth and Skurray, 1992) are some of the proteins involved in pilus synthesis and assembly (Achtman *et al.*, 1971, 1972). The TraA protein is pilin, subunits of which make the pilus, and are thought to pool in the inner membrane after synthesis (Moore *et al.*, 1981). The variation seen in pilin subunits from various plasmids is thought to provide a mechanism for plasmid specificity (Willetts and Maule, 1986). TraL, E, K, C and G are hydrophobic and are thought to be associated with the inner membrane, or found in the periplasm (Frost *et al.*, 1994). TraS and TraT have been assigned the role of entry and surface exclusion, respectively (Kingsman and Willetts, 1978). TraS is thought to be associated with the inner membrane and TraT with the outer membrane. Cells containing these proteins are found to block both mating pair formation and DNA transfer (Achtman *et al.*, 1977). TraG and TraN are responsible for mating pair stabilization during conjugation. TraN has been localized to the outer membrane (Maneewannakul *et al.*, 1992) and has been suggested to interact with OmpA and LPS moieties in the recipient cell (Klimke and Frost, 1998). Deletion of *traN* results in a greatly reduced transfer efficiency. TraG has been localized to the inner membrane and only the carboxyl region is required for mating pair stabilization (Manning *et al.*, 1981; Firth and Skurray, 1992). TraJ, Y, FinO and the antisense RNA, FinP, are involved in gene regulation. TraJ is a positive regulator of the P_Y promoter, which transcribes the majority of the *tra* genes in the F plasmid (Willetts, 1977). TraY is thought to be involved with the control of P_Y promoter as well, since F *traY* mutants are bacteriophage resistant and have decreased numbers of pili (Maneewannakul *et al.*, 1996). However, TraY from R100 acts as an autoregulator, negatively controlling transcription from P_Y (Taki *et al.*, 1998). FinP is a 75 base antisense RNA which duplexes with the untranslated region of the *traJ* transcript. Upon duplex formation, the double-stranded RNA is quickly degraded. The FinO protein prolongs the half-life of FinP (Lee *et al.*, 1992) and promotes duplex formation between FinP and the *traJ* transcript (van Biesen and Frost, 1994). R100-1 is a derepressed mutant of R100, and contains an insertion of one nucleotide in the *finO* gene making it non-functional (Yoshioka *et al.*, 1987)

Tra proteins involved in DNA metabolism: TraY

Proteins required for F plasmid DNA metabolism are TraY, TraI, TraM and TraD. TraY is required as an accessory protein allowing TraI to cleave at *nic* (Nelson *et al.*, 1995; Howard *et al.*, 1995). TraY binds to three sites in *oriT* (Lahue and Matson, 1990; Nelson *et al.*, 1993; Luo *et al.*, 1994): two near *nic* (*sbyA* and *sbyC*; Figure 1.2a), and another at the P_Y promoter (*sbyB*). Binding occurs as a monomer or a dimer (Nelson and Matson, 1996) and binding to *sbyA* induces DNA bending by approximately 50° (Luo *et al.*, 1994). R100 TraY is similar to F TraY, binding to a similar site in the *oriT* of R100 (*sbyA*; Figure 1.2b) and to two sites near the P_Y promoter (*sbyB* and *sbyC*; Inamoto and Ohtsubo, 1990). It has been suggested that F *traY* has undergone a gene duplication event (Inamoto *et al.*, 1988; Taki *et al.*, 1998). F TraY has now been classified into the Arc and Mnt repressor family because it contains a ribbon-helix-helix motif involved in dimerization and DNA binding (Schildbach *et al.*, 1998).

TraI

TraI was originally named DNA helicase I. Abdel-Monem *et al.* found in 1983 that this helicase came from the *traI* gene in an Hfr strain of *E. coli* which had the F plasmid integrated into the chromosome (Abdel-Monem *et al.*, 1983). TraI is a bifunctional protein that cleaves the *nic* site as a relaxase and then progressively unwinds the double-stranded DNA as a helicase (Traxler and Minkley, 1988). The relaxase and helicase domains are in the amino- and carboxyl-domains of TraI, respectively. The carboxyl-termini of relaxases are very divergent and may allow for the interaction with other plasmid specific proteins. Cleavage is site- and strand-specific (Matson and Morton, 1991) and results in the covalent linkage of the protein to the 5' end of the DNA (Matson *et al.*, 1993) leaving the 3' end of the DNA free. Attempts to determine if this is the site where donor strand DNA replication begins have been inconclusive (reviewed in Lanka and Wilkins, 1995). Cleavage of the DNA occurs as a phosphodiester transfer reaction, and the DNA is most likely covalently connected to the protein through a tyrosine residue. Tyrosine at position 23 was suggested to be the active tyrosine in F TraI (Byrd and Matson, 1997), however, tyrosine residues are also present at positions 16, 17, and 24. Sequence similarity to TrwC of plasmid R388 where the active tyrosine has been identified (Grandoso *et al.*, 2000) suggests that tyrosine at position 16 may be responsible for cleavage. Similar cleavage mechanisms are proposed for other relaxases such as TraI from RP4 (Pansegrau *et al.*, 1990), TraI from R100 (Inamoto *et al.*, 1991), TrwC from R388 (Llosa, *et al.*, 1995), and MobA from RSF1010 (Scherzinger *et al.*, 1992). TraI is mechanistically similar to the gene A protein from bacteriophage ϕX174 (Baas and Jansz, 1988; Waters and Guiney, 1993) and Rep proteins of gram-positive rolling circle replicating plasmids such as RepC from pT181, RepA from pC194-type plasmids, and RepB from pE194-type plasmids (Lanka and Wilkins, 1995; Pasengrau *et al.*, 1994). These relaxase and Rep proteins also share a common domain called the HUH motif ($HUHU_3$; H-histidine, U-bulky hydrophobic residue), which may coordinate metal ions needed for cleavage (Koonin and Ilyina, 1993). Nicking is performed by covalent linkage of the protein to the 5' end of the DNA through an active tyrosine, or hydrolysis of the origin using glutamic acid or aspartic acid. In either case, the free 3' end of the DNA is used to prime replication of the plasmid. One difference between the RCR Rep proteins and the conjugational relaxases is that RNA priming may be required for conjugal donor DNA synthesis in the F plasmid (Kingsman and Willetts, 1978). Another difference is that the relaxases noncovalently bind to the 3' side of the cleavage site and maintain plasmid superhelicity, perhaps waiting for a mating signal (Zechner *et al.*, 2000) while the Rep proteins release the 3' end allowing access to the polymerases. A third difference is that the active site tyrosine in relaxases is usually found within the first 50 amino acids of the protein, whereas Rep proteins contain this active site in the carboxyl-terminus of the protein. The F TraI cleavage/religation reaction at *nic* is likely a continuous process as long as all factors are present (Sherman and Matson, 1994). Previous cleavage assays showed that only TraI and supercoiled DNA were required for nicking (Matson and Morton, 1991), however, the presence of TraY and IHF (integration host factor) stimulated cleavage at *nic* of supercoiled DNA and promoted cleavage of linear DNA substrates containing *nic* (Nelson *et al.*, 1995). Cleavage is most efficient when TraY and IHF are added before TraI (Howard *et al.*, 1995). IHF is a host-encoded protein which binds to the *oriT* regions of F (Tsai *et al.*, 1990), pED208 (another IncF plasmid; Di Laurenzio *et al.*, 1995), R100 (Dempsey and Fee, 1990), and R388 (Moncalian *et al.*, 1999). IHF bends DNA at angles up to 140° and is thought to provide some type of quaternary structure to these sites (Thompson and Landy, 1988). The term relaxosome has come to refer to the relaxase and plasmid- and/or host-encoded factors which form a stable nucleoprotein complex at *oriT* (Zechner *et al.*, 2000).

Relaxases

Relaxases have been grouped based on sequence similarity at their active sites (Zechner *et al.*, 2000). The IncP-type relaxases, such as TraI from RP4, have a variable carboxyl-domain, and three functional motifs. Motif I is the catalytic domain and TraI from RP4 contains a tyrosine at position 22 which is responsible for cleavage (Balzer *et al.*, 1994). Since only one tyrosine is present it is also thought that RP4 TraI may function as a dimer (Pansegrau and Lanka, 1996b). An equilibrium between covalently closed plasmid and cleaved plasmid is thought to occur, which explains why 100% cleavage is never observed (Pansegrau and Lanka, 1996a). Motif II of RP4 TraI is thought to bind to the 3' end of the *nic* site, preventing loss of plasmid superhelicity (Pansegrau *et al.*, 1994). Motif III is thought to contribute to the catalytic activity of the tyrosine and contains the HUH motif (Pansegrau *et al.*, 1994). RP4 TraI requires TraJ, a homodimeric specificity determinant, for cleavage (Ziegelin *et al.*, 1989). TraJ is thought to bind to *oriT* first and alter the local DNA structure, allowing access for TraI (Pansegrau *et al.*, 1990). TraH is also required *in vivo* and interacts with TraI and TraJ to stabilize the relaxosome. TraK, also required *in vivo* (Furste *et al.*, 1989), wraps a 180 base pair region around itself, which may alter the superhelical density at *nic* (Ziegelin *et al.*, 1992). IncF- (F TraI) and IncW- (R388 TrwC) type relaxases also have the relaxase activity localized to the amino-terminus, and the helicase function localized to the carboxyl-terminus (Byrd and Matson, 1997; Llosa *et al.*, 1996). Cleavage is similar to the IncP-type, however, two tyrosines are thought to be involved (Zechner *et al.*, 2000). Mutational analysis of TrwC has suggested that tyrosine at position 18 is responsible for the initial cleavage and that tyrosine at position 26 performs the second strand-transfer reaction responsible for termination (Grandoso *et al.*, 2000; corresponding to tyrosines at position 16 and 23 in F TraI respectively) in a mechanism similar to the "flip-flop" mechanism of ϕX174 gene A protein (Hanai and Wang, 1993). Both TraY and IHF are required by TraI in F and R100 (another IncF plasmid) for nicking (Nelson *et al.*, 1995; Inamoto *et al.*, 1994). No direct evidence for interaction between these three proteins has been shown. However, since both TraY and IHF bend DNA, they are thought to have a conformational effect on the DNA which may be required for nicking (Luo *et al.*, 1994; Thompson and Landy, 1988). This conformational effect may also facilitate the required protein:protein interaction

(Byrd and Matson, 1997). It has also been proposed that this conformational effect may cause unwinding at the *nic* site allowing efficient cleavage (Nelson *et al.*, 1995). TraM has also been shown to bind to *oriT* but has no noticeable effect on cleavage *in vitro*, however, increased cleavage has been noted in the presence of TraM *in vivo* in the plasmid R1 (another IncF plasmid; Kupelwieser et al., 1998). The IncQ-type relaxases (such as MobA from RSF1010) contain an amino-terminal relaxase and a carboxyl-terminal primase (Scherzinger *et al.*, 1992). Cleavage is similar to the IncP-type relaxases using tyrosine at position 24 for cleavage (Scherzinger *et al.*, 1993). Cleavage *in vivo* is enhanced by the accessory proteins MobB and MobC. MobC is thought to increase the size of the unwound region near *nic* (Zhang and Meyer, 1997) and MobB is thought to increase the stability of the nicked species (Perwez and Meyer, 1996).

TraD

F TraD mutants are able to synthesize pili (Achtman *et al.*, 1972) and are able to form stable mating pairs. However, they are unable to effect DNA transfer (Manning *et al.*, 1981). Conjugal donor cell DNA synthesis also occurs in *traD* mutants, suggesting that TraD functions after this point during conjugation (Manning *et al.*, 1981). F TraD is localized to the inner membrane (Panicker and Minkley, 1992) and has two membrane spanning regions, with the amino- and carboxyl-terminal regions of the protein in the cytoplasm (Lee *et al.*, 1999). F TraM (a DNA binding protein required for transfer) is localized to the cytoplasm, however, *in vitro* assays have shown that TraM and TraD interact (Disque-Kochem and Dreiseikelmann, 1997). During overexpression of TraD and TraI, TraI was found to be membrane-associated only in the presence of TraD, suggesting that these two proteins interact (Dash *et al.*, 1992). This is similar to the observed interactions between RP4 TraG (the F TraD homologue) and the Mob protein (homologous to F TraI) from the mobilizable plasmid pBHR1 (Szpirer *et al.*, 2000). These experiments suggest that TraD is a coupling protein between the membrane and *oriT* and suggest that TraD may localize *oriT* to the transfer apparatus by interacting with DNA-bound proteins. The carboxyl-terminal of F TraD is important for the mobilization of F-specific plasmids, and its removal increases its ability to mobilize other plasmids but decreases its ability to mobilize F-specific plasmids (Sastre *et al.*, 1998).

TraM

The gene encoding *traM* is located closest to *oriT* in the F, R100, and R1 plasmids (Figure 1.2). TraM proteins from these plasmids are 127 amino acids long and show a high degree of identity (Figure 1.3). F and R100 TraM are 89% identical, and R1 TraM is 78% identical to F and R100 TraM. F TraM was originally thought to be localized in the inner membrane (Achtman *et al.*, 1979), but was later found to be primarily cytoplasmic with small amounts in the inner membrane (Di Laurenzio *et al.*, 1992). F *traM* mutants (JCFL102, a F*lac traM* frameshift mutant) were able to form stable mating pairs (Achtman and Skurray, 1977; Manning *et al.*, 1981) and cleavage was observed (Willetts and Wilkins, 1984). However, transfer was reduced to background levels (Kingsman and Willetts, 1978). Donor cell conjugal DNA synthesis was also reduced in *traM* mutants suggesting that TraM functions in DNA metabolism (Manning *et al.*, 1981). TraM binds to the *oriT* region of F and ColE1 (an F mobilizable plasmid; Di Laurenzio *et al.*, 1992), pED208 (Di Laurenzio *et al.*, 1995), R100 (Abo *et al.*, 1991), and

Figure 1.3. Protein sequence comparison of TraM from F, R100-1, and R1. Consensus sequences of all three proteins is shown above each section and was done using Peptool (v.1.1).

```
Consensus  MA-V--Y-S-----K-N-I-E-RR-EGA---D-S-  35
F TraM     MAKVNLYISNDAYEKINAIIEKRRQEGAREKDVSF  35
R100 TraM  MARVILYISNDVYDKVNAIVEQRRQEGARDKDISV  35
R1 TraM    MAKVQAYVSDEIVYKINKIVERRRAEGAKSTDVSF  35

Consensus  S----MLLELGLRV-EAQMERKESAFNQ-EFNK-L  70
F TraM     SATASMLLELGLRVHEAQMERKESAFNQTEFNKLL  70
R100 TraM  SGTASMLLELGLRVYEAQMERKESAFNQTEFNKLL  70
R1 TraM    SSISTMLLELGLRVYEAQMERKESAFNQAEFNKVL  70

Consensus  LEC-VKTQS-VAKILGIESLSPHVSGN-KFEYANM  105
F TraM     LECVVKTQSSVAKILGIESLSPHVSGNSKFEYANM  105
R100 TraM  LECVVKTQSSVAKILGIESLSPHVSGNPKFEYANM  105
R1 TraM    LECAVKTQSTVAKILGIESLSPHVSGNPKFEYANM  105

Consensus  VEDIR-KVSSEMERFFP-ND-E  127
F TraM     VEDIREKVSSEMERFFPKNDDE  127
R100 TraM  VEDIREKVSSEMERFFPKNDEE  127
R1 TraM    VEDIRDKVSSEMERFFPENDEE  127
```

R1 (Schwab *et al*., 1991) in a plasmid or allelic specific manner (Kupelwieser *et al*., 1998; Fekete and Frost, 2000). However, the TraM binding site arrangements vary between plasmids. TraM also autoregulates transcription from its two promoters in F (Penfold *et al*., 1996), R1 (Schwab *et al*., 1993), and R100 (Abo and Ohtsubo, 1993). TraM is also thought to be involved in the regulatory network involved in *tra* gene expression. Although the actual mechanism is unknown, regulation has been suggested to occur by readthrough transcription from one of the *traM* promoters for R1 (Polzleitner *et al*., 1997), or from a putative promoter inside *traM* as suggested for R100 (Dempsey, 1994; Stockwell and Dempsey, 1997). These readthrough transcripts are thought to become targets for FinP, allowing *traJ* transcripts to be translated. The increase in TraJ levels would then activate transcription from the P_Y promoter. Readthrough transcripts have not been found in the F plasmid (Penfold *et al*., 1996) leaving the regulatory mechanism for F TraM unknown. R1 TraM forms stable tetramers in solution (Verdino *et al*., 1999) and the amino-terminal end of R1 TraM is important for DNA binding (Schwab *et al*., 1993). R1 TraM is required for cleavage at *nic* (Kupelwieser *et al*., 1998), however, this is not true for F TraM (Everett and Willetts, 1980). This key difference in function suggests that the two TraM proteins are functionally distinct, even though they share sequence identity (Figure 1.3).

Objectives

The objectives of this work were to purify and characterize the TraM protein. It was hoped that newer techniques would clarify some of the ambiguities from previous characterizations (Di Laurenzio, 1992; Penfold, 1995). In addition to biochemically characterizing the protein, domains in the protein responsible for multimerization were also to be addressed. Studies to identify these regions were done to explain the characteristics of TraM in solution and also to try and build a model of how TraM binds to its binding sites. Since TraM has three binding sites in the *oriT* region, it was also important to characterize the binding of the protein to these sites individually and in various combinations. Analyzing the binding properties of each site would also aid in the design of a model for TraM binding to *oriT*. Since the three TraM sites are not identical, it was hoped that a more rigorous characterization of the bases which are bound and protected by the protein would aid in the explanation of how binding occurs to each of these sites. It was also hoped that experiments could be designed to address the allele-specific nature of the TraM and TraI proteins which had been previously identified (Reeves and Willetts, 1974). These experiments were done to determine whether the allele-specific nature was based on DNA binding or based on protein:protein interactions occurring at *oriT*.

Chapter 2

Materials and Methods

Materials and Methods

Bacterial strains, plasmids, and growth conditions. The following *Escherichia coli* strains were used in this study: CS2198 [*waaJ19*::Tn*lacZ* of CS1999] (Pradel *et al*., 1992); DH5α [Δ*lacU169* (φ80 *lacZΔM15*) *supE44 hsdR*17 *recA1 endA1 gyrA96* (Nal^r) *thi-1 relA1*] (Ausubel *et al*, 1987; Hanahan, 1983); ED24 [F$^-$ Spc^R Lac$^-$] (Willetts and Finnegan, 1970); ED2149 [F$^-$ *lacΔU124* Δ(*nadA aroG gal attL bio*)] (Dempsey and Willetts, 1976); JE2571-1 (Kingsman and Willetts, 1978); and XK1200 [F$^-$ Nal^R *lacΔU124* Δ(*nadA aroG gal attL bio gyrA*)] (Moore *et al*., 1987). Cells were grown in Luria-Bertani (LB) medium (Ausubel *et al*., 1987) or on LB with 1.5% agar (Difco Laboratories) supplemented with the appropriate antibiotics at the following final concentrations: 50 μg/ml ampicillin, 25 μg/ml kanamycin, 10 μg/ml tetracycline. *E. coli* strains carrying plasmid constructs were picked from selective plates, inoculated into 3 mL LB containing the appropriate antibiotics and grown overnight at 37°C (except where noted) on a Roller-Drum (Bellco) at medium speed.

Recombinant DNA techniques. Restriction enzymes, alkaline phosphatase (AP), and T4 DNA ligase were supplied by Roche and used following standard procedures (Ausubel *et al*., 1987) except as noted. Plasmids were transformed using CaCl$_2$ competent cells (Sambrook *et al*., 1989) or by electroporation using a Bio-Rad Gene Pulser at 2.5 V, 25 μFD, and 200 Ω. DNA fragments used to create plasmid constructs were isolated from acrylamide or agarose gels. Isolation from acrylamide was done by crushing the excised bands containing the fragments and eluting overnight in 300 μl of

500 mM ammonium acetate and 1 mM EDTA at 37°C, followed by phenol extraction and ethanol precipitation. Isolation from agarose was done using the Qiagen Gel Extraction Kit. PCR reactions were done using Vent polymerase (New England Biolabs) with 20 mmoles dNTP's (Roche), approximately 500 pmol of each primer in a 100 µL volume; cycled 30 times. Fill in reactions were done using Klenow polymerase (Roche) and 500 pmol dNTP in 30 µL at 37°C for 30 minutes. Gels placed on a Phosphor Screen (Molecular Dynamics) were analysed by a Molecular Dynamics Phosphor Imager 445SI using Image QuaNT version 4.2a software. Plasmids were isolated using the method of Birnboim and Doly (Birnboim and Doly, 1979) or using Qiagen Miniprep columns. .

Construction of chimeric plasmids. pNY300 (Frost *et al.*, 1989) was constructed by digesting F plasmid DNA with *Bgl*II and inserting the 1080 bp fragment into the *Bam*HI site of pUC18 (USB). pRF105 was constructed using a serendipitous mutation in R100-1 which created a *Bam*HI site 135 bp upstream of *nic* (Frost *et al.*, 1994). Digestion with *Bam*HI generated a 1045 bp fragment which was inserted into the *Bam*HI site of pUC18. pRF315 was constructed by digesting a PCR product generated from pRF105 using LFR51 (AAATAGAGAGTCGTTGGCGATCC) and Reverse (TCACACAGGAAACAGCTATGACCA) primers with *Eco*RI to give an 830 bp fragment. This was ligated to the 260 bp *Dra*I and *Hin*dIII fragment of pNY300 and inserted into pUC18 digested with *Eco*RI and *Hin*dIII. pRF206 was constructed by digesting a PCR product generated from pNY300 using the Universal (GGGTTTTCCCAGTCACGACG) and RFE4 (AAAACGTAAATCAGCAAAAACTTGTT) primers with *Hin*dIII to give a 209 bp fragment. This was ligated to an 888 bp fragment of pRF105 digested with *Eco*RI and *Dra*I and inserted into pUC18 digested with *Eco*RI and *Hin*dIII. pKJ4 is an *Eco*RV - *Eco*RI fragment containing *traY* and *traA* cloned into pT7.4 (same as pT7.3 except β-lactamase is in reverse orientation; Tabor and Richardson, 1985). This construct was created using the *Eco*RV site in *traJ* (Frost *et al.*, 1994) and an *Eco*RI site engineered by PCR into the 3' end of *traA*. All plasmids were sequenced using Sequenase (USB) and [α^{33}P]ddNTP (Amersham) to verify their construction.

Plasmid nicking assays. Nicking assays involving pOX38-Km and its derivatives were done as previously described (Frost and Manchak, 1998; Perwez and Meyer, 1996). For chimeric plasmids 3 ml cultures containing a chimeric plasmid were grown to OD_{600} of 0.4. Cells were lysed and plasmid DNA purified using the complete method of Birnboim and Doly (Birnboim and Doly, 1979). DNA was dissolved in 30 μl of Milli-Q® water. 2 μl of this DNA was completely digested with the appropriate restriction enzyme. The DNA was ethanol precipitated and dissolved in 10 μl of Milli-Q® water. Of this, 0.1 or 0.01 μl was added to the nicking reaction depending on DNA concentration. The nicking reaction mixture was 11.5 μl of: 41.5 μl Milli-Q® water, 5 μl of 10X Thermopol Buffer, 1 μl of 10 mM dNTP, 500 pmol of Universal primer, and 2 μl (approximately 20 μCi) of [α-^{32}P]dCTP (Amersham Pharmacia Biotech). Reactions were denatured for 2 minutes at 94°C before the addition of 2 μl of diluted Vent Polymerase (0.5 μl polymerase with 8 μl Milli-Q® water) (New England Biolabs). Reactions were thermocycled using a MJ Research MiniCycler at 94°C for 30 seconds, 61°C for 30 seconds, and 72°C for 1 minute for 35 cycles. The reactions were then removed, rolled on parafilm to remove remaining mineral oil and ethanol precipitated. DNA was dissolved in 15 µl Milli-Q® water and 5 µl of Sequencing Stop Solution (USB). 10 µl of each reaction was then denatured at 85°C for 10 minutes and loaded onto a 6% polyacrylamide gel containing 8 M urea. A dideoxy-sequencing reaction of each plasmid using Universal primer was performed with Sequenase (USB) and was loaded as a standard.

Quantitation of nicking efficiency. Gels containing the nicking and sequencing reactions were exposed to a Molecular Dynamics Phosphor Screen overnight. Bands located at *nic* were compared to bands located at the *Dra*I or *Hin*fI restriction enzyme sites to determine the percent of cleavage in each sample. *Dra*I was used as the restriction enzyme for pNY300 and pRF315, while *Hin*fI was used for pRF206 and pRF105. Occasionally other prominent bands were also found in a single lane and the values of these bands were added to those of the bands located at the restriction enzyme sites. Background values were also subtracted from both band intensities at *nic* and the restriction enzyme sites.

Mobilization efficiency assays. Recipient and donor cells were grown to early log phase (OD_{600} of 0.4) with appropriate antibiotic selection. Cells were washed twice and resuspended in the same volume of medium. 100 µl each of donor and recipient cells were added to 800 µl of medium and incubated at 37°C for 30 minutes. Cells were vortexed and placed on ice. Serial dilutions of the mating cultures were made using 1X SSC (0.15 M sodium chloride, 0.015 M sodium citrate, pH 7.0). 10 µl of each dilution were spot-dropped onto selective plates containing combinations of antibiotics to select for transconjugants containing mobilizable plasmids or self-transmissible plasmids, donors, or recipients. Plates were dried and incubated at 37°C overnight. Mating efficiency is reported as number of transconjugants per 100 donors.

Cell fractionation. Cells were fractionated using a modification of previously published methods (Di Laurenzio *et al.*, 1991; Noltman *et al.*, 1961). F*lac* derivatives were introduced into *E. coli* ED2149, while all other plasmids were in DH5α. 200 ml of cells were grown to OD_{600} of 0.8 at 37°C. Cells were cooled on ice and pelleted in a Sorvall Superspeed RC2-B centrifuge at 10,000 X *g* in a GSA rotor for 10 minutes. All steps after this point were carried out at 4°C with chilled solutions. Cells were washed with 30 ml of 50 mM Tris-HCl, pH 8 and 25% sucrose and then pelleted. Cells were resuspended in 5 mL of Milli-Q® water containing 40 μg/ml of DNase I and RNase A and lysed by three passages through a French Pressure Cell Press (American Instrument Co.) at 13,000 psi. Unbroken cells were pelleted at 1,200 X *g* in an SS-34 rotor for 20 minutes. The supernatant was removed and represented the total cell lysate. Membranes were pelleted from the supernatant at 100,000 X *g* in a SW-41 rotor (Beckman Coulter Corp.) for 1 hour. Pellets were resuspended in 3 ml of 10% sucrose and 5 mM EDTA and fractions constituting unwashed membranes were removed and frozen. Membranes were pelleted and resuspended in 2 mL of 10% sucrose and 5 mM EDTA to give the final washed membrane fraction. Protein concentrations were determined using the Lowry assay (Lowry *et al.*, 1951) and glucose-6-phosphate dehydrogenase assays were performed to determine contamination of membranes with cytoplasmic proteins. For the glucose-6-phosphate dehydrogenase assays approximately 50 μl of sample (1/60 of total membrane fractions) was added to 1 ml of 83 mM glycylglycine, 1 mM glucose-6-phosphate, 0.3 mM NADP, and 10 mM magnesium sulphate. The increase in A_{340} was measured and the molar extinction coefficient of NADPH (6.2×10^3) and protein concentrations were used to determine the specific activity for each fraction (Noltman *et al.*, 1961).

Immunoblot analysis. Samples were run on 15% or 8% SDS polyacrylamide gels with 7% stacking gels at 35 mA with a whole cell sample and purified TraM (Di Laurenzio *et al.*, 1992) as controls. Whole cell samples were made by pelleting 0.1 $O.D._{600}$ of cells at 4,000 X *g* in a Eppendorf MicroCentrifuge (Model 5415C), resuspending in 1X Loading Dye (Sambrook *et al.*, 1989) and boiling for 10 minutes prior to loading. Proteins were transferred to a PVDF membrane (Millipore Immobilon-P) using a Bio-Rad Trans-Blot Electrophoretic Transfer Cell at 250 mA for 2 hours at 4°C. Membranes were blocked overnight in 10% skim milk (Gibco) in 1X TBST (50 mM Tris, 150 mM sodium chloride, 0.1% Tween 20 (v/v; Caledon Laboratories) at 4°C. Membranes were incubated at room temperature with anti-TraM (1/25 000; Di Laurenzio *et al.*, 1992), anti-TraD (1/20 000) or anti-TraI (1/20 000; kind gifts of K. Ippen-Ihler) and donkey anti-rabbit secondary antibodies linked to horseradish peroxidase (1/5 000; Amersham Life Science) for one hour each. Detection was performed with Renaissance Western Blot Chemiluminescence Reagent (NEN Life Science Products) and exposed on X-OMAT AR film (Kodak).

Construction of pRFM2T75. The gene encoding TraM was isolated from pNY300 by digesting with *Bst*BI and blunting the ends with Klenow polymerase. The DNA was then digested with *Sac*I and the 665 bp fragment purified from a 1.2 % agarose gel. This fragment was ligated into pT7-5 (same as pT7.3 except β-lactamase is in reverse orientation; Tabor and Richardson, 1985) which had been digested with *Eco*RI, blunted using Klenow polymerase, then digested with *Sac*I and purified from a 0.6 % agarose gel. Clones were isolated using Qiagen mini prep columns and confirmed by digesting with *Sal*I and subsequent sequencing. Nomenclature for plasmid naming was RF (Richard Fekete) M (TraM) 2 (second attempt at cloning) T75 (placed into vector pT7.5).

Construction of His-tagged TraM clones. pRF400 and pRF401 were constructed by using PCR and primers RFE6 (GGATCCATGGCTAAGGTGAACCTG) and LFR22 (TCACACAGGAAACAGCTATGACCA) to amplify *traM* from pSPE2309 (wild-type TraM; Penfold, 1995) and pSPE2307 (C-terminal 8 amino acid deletion; Penfold, 1995) respectively. DNA was digested with *Bam*HI and *Kpn*I and cloned into pQE40 (Qiagen) digested with the same enzymes. pRF402 was created due to a PCR incorporation error which mutated isoleucine 109 to threonine.

Overexpression and purification of TraM. pRFM2T75 was transformed into DH5α already containing plasmid pGP1-2 (Tabor and Richardson, 1985) a plasmid containing the T7 RNA polymerase gene. A single colony was used to inoculate 2 mL of LB containing kanamycin and ampicillin and grown for 6 hours at 30°C. This was then subcultured into 20 mL LB (containing kanamycin and ampicillin) and grown at 30°C overnight. This was then subcultured into 1 L of LB (containing kanamycin and ampicillin) and grown at 30°C until the culture reached an $O.D._{600}$ of 1.0. The flask was placed under running hot water for 1 minute to quickly heat the culture, then incubated at

$42^{\circ}C$ for 30 minutes followed by 45 minutes at $37^{\circ}C$. Cultures were pelleted at 4000 rpm for 15 minutes in a GSA rotor. 0.1 O.D.$_{600}$ of cells was run on 18 % SDS polyacrylamide gels to determine expression levels. Cells were resuspended in 6 mL of chilled 50 mM Tris pH 8, and 25 % sucrose. Only chilled solutions were used after this step. Cells were pelleted at 7000 rpm for 5 minutes in a GSA rotor and resuspended in 5 mL of 50 mM Tris pH 8, and 25 % sucrose and 35 μL of lysozyme (12 mg/mL in 0.25 M EDTA) was added. The cells were then passed through a French Pressure Cell Press (American Instrument Co.) three times to ensure lysis and 150 μL of RNase A (5 mg/mL) was added. The lysate was centrifuged at 100 000 X g in a SW-41 rotor for 1 hour to pellet the membranes. The membrane pellet was resuspended in 500 μL of 50 mM Tris, 0.1 M EDTA, and 1 mM DTT. To the supernatant (14 mL) 3.5 g of ammonium sulfate was slowly added to eliminate high local concentrations of salt, and the mixture was placed on a rocker at $4^{\circ}C$ for 30 minutes. The mixture was centrifuged at 10 000 X g (9 500 rpm in a SS34 rotor) for 30 minutes. The pellet was dissolved in 2 mL 50 mM Tris, 0.1 M EDTA, and 1mM DTT. The supernatant and pellet were desalted using a PD-10 desalting column (Pharmacia) and eluted in 2.5 mL 50 mM Tris. Protein concentrations were determined using a Lowry assay (Ausubel et al., 1987) and approximately 20 μg was run on a 18% SDS polyacrylamide gel to determine the fractions containing TraM.

Column chromatography. 15 mg of crude TraM were loaded onto a MonoQ HR 5/5 column (Pharmacia) using a Pharmacia FPLC model LCC-500 and washed with 50 mM Tris pH 8. A salt gradient (mixing 50 mM Tris pH 8 and 1 M sodium chloride in 50 mM Tris pH 8) was used to elute the bound protein, which eluted at a salt concentration of approximately 250 mM sodium chloride. Lowry assays (Ausubel *et al.*, 1987) were performed on all fractions and approximately 20 µg of each were run on an 18% SDS polyacrylamide gel. Fractions containing TraM were pooled and desalted using a PD-10 desalting column and eluted in 2.5 mL 50 mM Tris pH 8. Approximately 1 mg of protein was then loaded onto a 1 mL Native DNA Cellulose column (Amersham Pharmacia Biotech) equilibrated with TED buffer (100 mM Tris, 0.1 mM EDTA, 1 mM DTT). The column was washed with TED at a flow rate of 2 mL/minute at 4^{O}C and the protein was eluted with 0.1, 0.2, 0.3, 0.4 M sodium chloride in TED for 1 hour at each salt concentration. TraM eluted at a concentration of approximately 300 mM sodium chloride. Fractions containing TraM were combined, then concentrated using an Amicon Stir Cell with a YM3 membrane (Diaflo). Half of the concentrate was desalted using a PD-10 column and eluted with TED. The other half was desalted and eluted with 20 mM potassium phosphate for analytical purposes. Both were concentrated to approximately 1 mg/mL using the Amicon Stir Cell.

Protein analysis. Mass spectrometry was performed on TraM in 20 mM potassium phosphate using Fisons VG Quattro Electrospray Mass Spectrometer collecting masses between 12 000 and 17 000 da. Amino acid composition was determined using a Beckman 6300 Amino Acid Analyzer. Samples were hydrolyzed under vaccum in 6 M hydrochloric acid, 0.1 % phenol at 160^{O}C for 1 hour and passed over a 12 cm ion-exchange column. Amino acids were eluted with ninhydrin and detected in a flow cell using absorbance at 570 nm. Amino acid identity was determined by elution time from the ion-exchange column. Amino terminal sequencing was performed using a Hewlett

Packard G1005 sequencer using Edman degradation and phenylisothiocyanate to generate phenylthiohydantoin-amino acid derivatives. The PTH-amino acid derivatives were separated by Reverse Phase HPLC and detected by absorbance at 269nm. Identity of each amino acid was determined by retention time on the column in comparison to a set of PTH-amino acid standards. All analysis was performed at Alberta Peptide Institute.

Size Quantification of TraM. Analytical ultracentrifugation (XL-1 Analytical Ultracentrifuge, Beckman) was performed on TraM in potassium phosphate and diluted in potassium chloride or potassium phosphate. Experiments were performed at 3 different concentrations (3, 7, and 13 μM), at 2 different centrifugal speeds (12 000 and 16 000 rpm), and in 100 mM, 300 mM potassium chloride or 20 mM potassium phosphate. Data were plotted as concentration (AU) vs. (radial distance)^2/2 of the sample cell and mathematical equations for various species were fit to the data. These plots were then analysed using residual plots and using the square root of the variances (less than 2×10^{-2}) to determine the fit for each species. Size exclusion chromatography was performed on 400 μg of purified TraM in 20 mM potassium phosphate on a Superose 12 column (Pharmacia) using a Pharmacia FPLC model LCC-500. Elution was performed using 20 mM potassium phosphate, 100 mM potassium chloride. As standards, 250 μg of each of α-lactalbumin, carbonic anhydrase, chicken egg albumin, and bovine serum albumin (Sigma) were used.

Construction of TraM binding site clones. pRF911 was constructed by phosphorylating approximately 400 pmol of RFE11 (CTAGGGACGCACCGCTAGCAGCGCCCCTAGCGGTATC) and RFE12 (CTAGGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCC) with T4 polynucleotide kinase (Roche), heating to 95°C, cooling to room temperature (to anneal primers), digesting with *Bln*I for 30 minutes, and ethanol precipitating the DNA. This was ligated to 0.3 pmol of pBEND2 (Kim *et al*, 1989) digested with *Xba*I, dephosphorylated with alkaline phosphatase (Roche), and purified from a 0.6% agarose gel. pRF912 was a construction error where 2 *sbmA* sites were cloned in tandem into pBEND2. pRF918 was constructed by annealing 400 pmol of LFR49 (CTAGAGCAGCGCCCCTAGCGG) and LFR50 (CTAGCCGCTAGGGGCGCTGCT) and ligating to 0.3 pmol of pBEND2 digested with *Xho*I. pRF920 was constructed by digesting pNY300 with *Rsa*I and *Dra*I and purifying the 58 bp band from an agarose gel. This was ligated to pBEND2 digested with *Sal*I, dephosphorylated with alkaline phosphatase, and filled in with Klenow polymerase. pRF940 was constructed by digesting pNY300 with *Dra*I and *Sal*I, and purifying the 300 bp fragment from a 5% polyacrylamide gel. This was then digested with *Bst*BI and filled in with Klenow polymerase and the 190 bp fragment was purified from a 8% acrylamide gel. This was ligated to pBEND2 digested with *Xba*I, filled in with Klenow polymerase, dephosphorylated with alkaline phosphatase and isolated from a 0.6% agarose gel. pRF930 was constructed by digesting pRF940 with *Bam*HI, *Cla*I, and *Rsa*I, filling in with Klenow polymerase, and purifying the 144 bp fragment from an 8% acrylamide gel. This was ligated to pBEND2 digested with *Xba*I, dephosphorylated with alkaline phosphatase, filled in with Klenow polymerase and isolated from a 0.6% agarose gel.

Chemical Crosslinking. PCR using RFE2 (GGTGCCTGACTGCGTTAGCA) and RFE3 (TAGGCGTATCACGAGGCCCT) was used to amplify *sbmA* (pRF911), half *sbmA* (pRF918), *sbmC* (pRF920), *sbmAB* (pRF930), and *sbmABC* (pRF940). DNA was used at a concentration of 25 nM and TraM was used at a concentration of 150 nM in binding reactions containing 20 mM sodium phosphate, 150 mM sodium chloride in a volume of 15 μL. Glutaraldehyde (Sigma) was used at final concentrations of 0.00025% to 25%. Dithiobis(succinimidylpropionate) (DSP; Pierce) and Bis(sulfosuccinimidyl)suberate (BS^3; Pierce) were used at final concentrations of 100 ηM to 1mM in 20 mM sodium phosphate, 150 mM sodium chloride. Protein was allowed to bind to DNA at 37^{O}C for 15 minutes after which crosslinker was added and allowed to react at room temperature for various lengths of time. To stop BS^3 and DSP reactions, 1 μL of 1 M Tris pH 7.5 was added and incubated at room temperature for a further 15 minutes. 5 mL of 6X protein loading dye (350 mM Tris pH 6.8, 30% glycerol, 350 mM SDS, 600 mM DTT, 9 mM bromophenol blue) was added to all reactions and the reactions were placed at 100^{O}C for 10 minutes following which, samples were run on a 15% SDS polyacrylamide gel at 35 mA until the dye front reached the bottom of the gel. Proteins were transferred to Immobilon-P PVDF membrane (Millipore) at 250 mA for 2 hours in a Trans-Blot Electrophoretic Transfer Cell (Bio-Rad). Membranes were blocked overnight in 10% skim milk (Difco) in TBST at 4^{O}C. Primary (anti-TraM, 1/20 000 dilution; Di Laurenzio *et al.*, 1992) and donkey anti-rabbit IgG secondary antibodies linked to horseradish peroxidase (Amersham, 1/5,000) were incubated in 10% skim milk in TBST at room temperature, each followed by 4, 15 minute washes using TBST at room temperature. Detection was performed using Renaissance Chemiluminescence Reagent Plus (NEN) at room temperature for 5 minutes. Membranes were subsequently exposed to X-OMAT AR film (Kodak) for various times.

Determination of the oligomeric state of TraM bound to DNA. DNA templates containing *sbmA*, *sbmC*, *sbmAB*, and *sbmABC* were PCR amplified using the primers RFE9 (GCTGCCCGGGAGGCCTTC) and RFE10 (GCTGGATATCTTTAAACTCGAG), and radioactively labelled by including 5 μL (50 μCi) of $[\alpha^{32}\text{P}]$dCTP (Amersham) in the reaction. Reactions were dried down to 10 μL in a Concentrator (Savant) and run on a 1.5% agarose gel. Bands corresponding to *sbmA* (158 bp), *sbmC* (185 bp), *sbmAB* (266 bp), and *sbmABC* (319 bp) were excised from the gel and the DNA isolated using the Qiagen gel extraction kit. *sbmA* was digested with *Bam*HI and *Dra*I, *sbmC* was digested with *Rsa*I and *Bgl*II, *sbmAB* and *sbmABC* were digested with *Bam*HI and *Bgl*II. Digests were then run on an 8% polyacrylamide gel. The 97, 104, 167 and 220 bp fragments (respectively) were excised and isolated and resuspended in 100 μL Milli-Q® water. Specific activity was determined using a Scintillation counter LS3801 (Beckman), and DNA concentration was calculated (in pmol/μL) using the following equation:

$$\frac{\text{Spec. act. of PCR product (cpm/}\mu\text{L) X cold dCTP added to PCR (pmol)}}{\text{vol. }[\gamma^{32}\text{P]dCTP added (}\mu\text{L) X number of Cytosine/PCR product X Spec. act. of }[\gamma^{32}\text{P]dCTP (cpm/}\mu\text{L)}}$$

Approximately 2 fmol of DNA was used in each 15 μL binding reaction. Binding reactions were conducted by mixing 3 μL of 5X retardation buffer [250 mM Tris, 50% glycerol, 150 μg/mL bovine serum albumin (Roche), and 5 mM DTT], 1 μL of poly (dIdC), 1 mg/mL (Roche), 1 μL of DNA template, and 10 μL of TraM. TraM was used at final concentrations of 5 and 0.5 μM. Binding reactions were incubated at 37°C for 15 minutes and then loaded onto 7, 8, 9, and 10% gels which were prerun at 4°C for 20 minutes. Standards [15 μL α-lactalbumin, 20 μL carbonic anhydrase, 20 μL chicken egg albumin, 15 μL bovine serum albumin, and 25 μL urease (Sigma MW-ND-500)] were mixed with equal amounts of sample buffer [1 mL stacking gel buffer (5.98 g Tris and 460 μL TEMED diluted to 100 mL with water and pH adjusted to 6.7 with hydrochloric acid), 1 mL glycerol, 1 mL water, and 0.25 mg bromophenol blue]. Gels were run at 4°C at 35 mA using electrode buffer (1.2 g Tris and 5.76 g glycine dissolved in 2 L of water and pH to 8.3, stored at 4°C) until the dye was approximately 1 cm from the bottom of the gel. Gels were poured by mixing: 1) separating gel buffer (36.5 g Tris, 230 μL TEMED diluted to 100 mL with water and pH adjusted to 8.9 with 1 M hydrochloric acid), 2) separating acrylamide (28 g acrylamide, 740 mg N,N'-methylenebisacrylamide diluted to 100 mL with water), 3) sucrose (5 g in 100 mL of water), 4) freshly prepared ammonium persulfate (40 mg in 5 mL of water). After mixing, the solution was de-aerated for 30 seconds using a side arm flask and then poured into the gel apparatus to approximately 3 cm from the top. Water was added to the top of the gel and the gel was allowed to polymerize for 1 hour. To obtain various gel concentrations different amounts of each solution were added: 7%: 5.25, 10.5, 23.6, and 2.6 mL respectively; 8%: 5.25, 12, 22.1, and 2.6 mL respectively; 9%: 5.25, 13.5, 20.6, and 2.6 mL respectively; and 10%:

5.25, 15, 19, and 2.6 respectively. After polymerization the water was removed, the gel was rinsed with stacking buffer and the separating gel was poured. The separating gel consisted of 2 mL stacking buffer, 4 mL stacking acrylamide (10 g acrylamide and 2.5 g N,N'-methylenebisacrylamide dissolved in 100 mL water), 2 mL riboflavin (4 mg riboflavin in 100 mL of water), and 8 mL of sucrose solution. This mixture was de-aerated for 1 minute using a side arm flask and poured on top of the polymerized separating gel. Fluorescent light was used to polymerize the stacking gel for 1 hour. After electrophoresis, the bottom of the gel was cut off at the tracking dye marker (bromophenol blue) and fixed in fixative solution (40% methanol, 7% acetic acid, v/v) for 3 hours at room temperature. Gels were stained overnight in staining reagent (500 mg Coomassie brilliant blue R (Bio-Rad) in 500 mL fixative solution) at room temperature. Gels were destained in fixative solution with Kimwipes soaking up excess stain for 4 hours at room temperature. Gels were then fixed in 7% acetic acid for 2 hours. Gels were dried using a Gel Dryer 583 (Bio-Rad), migration distances of the protein markers were measured and compared to the distance run by the tracking dye, and the gels were exposed to a Molecular Dynamics Phosphor Screen overnight. The relative mobility (R_f) for each protein standard was calculated by measuring the distance travelled by each molecule divided by the distance migrated by the bromophenol blue marker. R_f for each molecule was placed into the equation: $100[\log(R_f \times 100)]$ and plotted versus gel percentage on normal graph paper. The negative slope was then plotted versus the molecular weight of each protein to create a standard curve. Unknowns were also plotted using $100[\log(R_f \times 100)]$ versus gel percentage, their negative slopes taken and molecular weight calculated from the standard curve.

Yeast two-hybrid system plasmid synthesis. All experiments used the yeast strain PJ69-4A (James *et al*., 1996). Wild-type TraM was cloned into pGAD-C1 and pGBD-C1 (James *et al*., 1996) from pRF400 using the *Bam*HI and *Pst*I sites (5' and 3' of *traM*, respectively) and called pRFAD127 and pRFBD127 respectively. The 6.6 kbp, 5.85 kbp, and 420 bp fragments were excised from 0.6% and 1.5% agarose gels. pRFAD24 was synthesized by digesting pRFAD127 with *Sal*I and isolating the 6726 bp fragment from a 0.6% agarose gel and religating the plasmid. pRFAD108 was synthesized by digesting pRF400 with *Bam*HI and *Eco*RV and isolating the 325 bp fragment from a 1.5% agarose gel. pGAD-C1 was digested with *Cla*I, filled in with Klenow enzyme, and digested with *Bam*HI and the 6.6 kbp plasmid was isolated from a 0.6% agarose gel. The 6.6 kbp fragment was ligated to the 325 bp fragment to make pRFAD108. pRFAD119 was synthesized by digesting pRF401 with *Bam*HI and *Pst*I and isolating the 395 bp fragment from a 1.5% agarose gel. pGAD-C1 was then digested with *Bam*HI and *Pst*I and the 6.6 kbp fragment was isolated from a 0.6% agarose gel. The 6.6 kbp fragment was ligated to the 395 bp fragment to make pRFAD119. pRFADI109T was synthesized by digesting pRF402 with *Bam*HI and *Pst*I and isolating the 419 bp fragment from a 1.5% agarose gel. pGAD-C1 digested with *Bam*HI and *Pst*I and the 6.6 kbp fragment was isolated from a 0.6% agarose gel. The 6.6 kbp fragment was ligated to the 419 bp fragment to give pRFADI109T. pRFADA37V was synthesized by using PCR to amplify the mutant TraM gene from pLFR23 (Penfold, 1995) using the RFE6 (GGATCCATGGCTAAGGTGAACCTG) and LFR22 (TCACACAGGAAACAGCTATGACCA) primers. The 600 bp PCR product was isolated from a 1% agarose gel. This was then digested with *Bam*HI and *Dra*I and the 510 bp product was isolated from a 1% agarose gel. pGAD-C1 was digested with *Cla*I, filled in with Klenow polymerase, then digested with *Bam*HI, and the 6.6 kbp fragment was isolated from a 0.6% agarose gel. The 6.6 kbp fragment was ligated to the 510 bp fragment to give pRFADA37V. pRFADC103 was synthesized by digesting pRFAD127 with *Hin*dII, *Pst*I, and *Bgl*II, and isolating the 346 bp fragment from a 1.5% agarose gel. pGAD-C1 was digested with *Cla*I, filled in with Klenow polymerase, digested with *Pst*I, and the 6.6 kbp fragment was isolated from a 0.6% agarose gel. The 6.6 kbp fragment was ligated to the 346 bp fragment to give pRFADC103. pRFADC19 was synthesized by digesting pRFAD127 with *Eco*RV and *Pst*I and isolating the 87 bp fragment from a 1.5% agarose gel. pGAD-C1 was digested with *Pst*I and *Sma*I and the 6.6 kbp fragment was isolated from a 0.6% agarose gel. The 6.6 kbp fragment was ligated to the 87 bp fragment to give pRFADC19. pRFBD108 and pRFBD24 were synthesized by digesting pRFAD108 and pRFAD24 with *Eco*RI and *Pst*I and isolating the 359 bp and 98 bp (respectively) fragments from a 1.5% agarose gel. These were ligated to pGBD-C1 digested with the same enzymes and purified from a 0.6% agarose gel. pRFBDC103 was synthesized by digesting pRFADC103 with *Bam*HI and *Pst*I and ligating the 354 bp fragment to pGBD-C1 digested with the same enzymes. pRFBDI109T and pRFBDC19 were synthesized digesting pRFADI109T and pRFADC19 with *Eco*RI and ligating the 408 and 75 bp fragments (respectively) to pGBD-C1 digested with the same enzyme and dephosphorylated with alkaline phosphatase.

Liquid media and plates. 10X Amino Acid mix was made by mixing uracil, L-arginine, L-methionine to a final concentration of 20 mg/mL, L-tyrosine, L-isoleucine, L-lysine to a final concentration of 30 mg/mL, L-phenylalanine to a final concentration of 50 mg/mL, L-valine to a final concentration of 150 mg/mL, and threonine to a final concentration of 200 mg/mL. This was filter sterilized and stored at 4^{O}C. 10X yeast nitrogen base was dissolved to a final concentration of 67 mg/mL, filter sterilized and stored at 4^{O}C. 40% glucose was autoclaved and stored at 4^{O}C. L-tryptophan stock solution was made to 1 g/L, filter sterilized and stored in a foil wrapped bottle at 4^{O}C. Adenine, L-histidine, and L-leucine stock solutions were made to 1 g/L, filter sterilized and stored at 4^{O}C. Each mL of liquid media contained: 100 µL 10X amino acid mix, 100 µL 10X yeast nitrogen base stock, 50 µL of 40% glucose, 20µL of each of adenine, L-tryptophan, L-histidine, and L-leucine stock, 670 µL water. Synthetic complete (SC) plates were made by mixing 70 mL agar (2.85 g/100mL), 10 mL 10X amino acid mix, 10 mL 10X yeast nitrogen base stock, 5 mL 40% glucose, 2 mL of adenine, L-tryptophan, L-histidine, and L-leucine stocks, and 200 µL of 1 M 3-aminotriazole (Sigma) if necessary. Plates were stored at room temperature until use. YPAD plates were made by mixing 10 g yeast extract (Bacto), 20 g peptone (Bacto), 50 mL 40% glucose, 100 mg adenine hemisulfate, and 20 g agar dissolved to 1 L with water and autoclaved.

Transformation of yeast. Two to three large colonies of yeast were picked and suspended in 1 mL of water in an Eppendorf tube. Cells were pelleted by spinning at 7000 rpm for 15 seconds in an Eppendorf model 5415C benchtop centrifuge. Cells were resuspended in 50 µL of 100 mM lithium acetate and incubated at 30^{O}C for 15 minutes.

Cells were repelleted and the liquid removed with a pipettor. To the cell pellet, 240 µL of 50% polyethylene glycol (PEG) in TE (10 mM Tris, 1 mM EDTA), 55 µL of 1 M lithium acetate, 25 µL of 1 mg/mL sonicated calf thymus DNA, and 30 µL of water containing approximately 1 µg of plasmid was added. Reagents were added in this order to protect the cells from the toxic effects of the concentrated lithium acetate. Cells were vortexed for 1 minute and then incubated at 30^{O}C for 30 minutes. Cells were heat shocked for 40 minutes at 42^{O}C, then 1 mL of SC media (without the appropriate amino acid) was added and the cells were placed at 30^{O}C for 1 hour. Cells were pelleted and resuspended in 1 mL of water. 200 µL was plated onto the appropriate selective plate and incubated at 30^{O}C for 3-4 days.

β-Galactosidase assays. Cells were grown in SC medium (without the appropriate amino acids) overnight at 30^{O}C to an OD_{600} of approximately 2. 200 µL of culture was added to 700 µL of Z buffer (16.1 g/L sodium phosphate, dibasic, 5.5 g/L sodium phosphate, 0.75 g/L potassium chloride, 0.246 g/L magnesium sulphate, pH 7) containing β-mercaptoethanol (0.27 mL/100mL Z buffer; BDH). 50 µL of chloroform and 50 µL of 0.1% SDS were added and vortexed for 30 seconds. 160 µL of ONPG (4 mg/mL in Z buffer; Sigma) was added, mixed, and incubated at 30^{O}C for 90 minutes. Reactions were quenched by adding 400 µL of 1 M sodium carbonate and the cell debris removed by centrifuging for 10 minutes at 14 000 rpm in a benchtop centrifuge (Eppendorf 5415C). Absorbance at 420 nm was read using a Spectramax Plus microplate reader, using Softmax Pro v2.4.1, correcting for lightpath length to 1 cm. β-galactosidase activity was calculated using the equation: $1000[OD_{420}/t \times V \times OD_{600}]$ where t is time in minutes, and V is volume of cells added to the reaction in mL.

Confirmation of mutants. Cells were grown overnight in 2 mL SC (without appropriate amino acids) at $30^{O}C$ and isolated using the Qiagen QIAprep protocol: cells were pelleted at 5000 X g for 5 minutes in a benchtop microfuge (Eppendorf 5415C) and resuspended in 250 µL of P1 buffer (Qiagen) containing RNase A (0.1 mg/mL, Roche). Approximately 100 µL of acid-washed glass beads (Sigma) were added and the mixture was vortexed for 5 minutes. The beads were allowed to settle and the liquid was removed from the top with a pipettor. 250 µL of P2 buffer (Qiagen) was added, mixed and let stand for 5 minutes at room temperature. 350 µL of N3 buffer (Qiagen) was added and centrifuged at 14 000 rpm for 10 minutes. The cleared lysate was passed through a QIAprep spin column by spinning at 14 000 rpm for 1 minute. The column was washed with 750 µL of PE buffer and DNA was eluted with 50 µL of water. 0.5 µL of this was used in a PCR reaction using RFE 15 (GTTGAAGTGAACTTGCGGGG) and either RFE13 (GGAAGAGAGTAGTAACAAAGG, specific for BD plasmids) or RFE14 (CTATTCGATGAAGATACCCC, specific for AD plasmids). 10 µL of each 100 µL reaction was run on a 1.8% agarose gel to determine the size of the products.

Electromobility shift assays. PCR was used to amplify sequences from pRF911, pRF912, pRF918, pRF920, pRF930, and pRF940 with primers RFE16 (GGTGCCTGACTGCGTTGCA) and RFE17 (TAGGCGTATCACGAGGCCCT). DNA was radio-labelled by including approximately 50 µCi $[\alpha^{32}P]$ATP in the PCR reactions.

Reactions were concentrated in a Speed-vac (Sorvall) and the DNA isolated from a 1.5% agarose gel. Radioactivity of [α^{32}P]ATP and labelled DNA was quantitated using the Scintillation counter LS3801 (Beckman). DNA concentration was calculated (in pmol/µL) using the equation:

$$\frac{\text{Spec. act. of PCR product (cpm/}\mu\text{L) X cold dCTP added to PCR (pmol)}}{\text{vol. [}\gamma^{32}\text{P]dCTP added (}\mu\text{L) X number of C/PCR product X Spec. act. of [}\gamma^{32}\text{P]dCTP (cpm/}\mu\text{L)}}$$

For single-stranded retardations, primers were end-labelled using [γ^{32}P]ATP and Polynucleotide Kinase, purified using Quick Spin Oligo Columns (Roche), and quantitated using the Scintillation counter. Approximately 3000 cpm of DNA was used in each retardation reaction along with the specified amount of purified TraM. Binding reactions were conducted in binding buffer (final concentration 50mM Tris, 10% glycerol, 1 mM DTT, 30 µg/mL) with 1µg poly (dIdC) in a final volume of 15 µL. After the addition of protein, binding reactions were incubated at 37$^{\mathrm{O}}$C for 15 minutes until loading on a 8% TB (89 mM Tris, 89 mM boric acid) acrylamide gel which had been prerun at 4$^{\mathrm{O}}$C at 30 mA. Gels were run at 4$^{\mathrm{O}}$C at 30 mA until the bromophenol blue marker reached the bottom of the gel. They were then dried and placed on a Phosphor Screen (Molecular Dynamics) overnight.

Hill plots. Hill plots were used in an attempt to quantify the cooperative DNA binding of TraM. Using the basic equilibrium equation for DNA binding:

$$nP + DNA \leftrightarrow Pn{\cdot}DNA$$

where n is the number of protein molecules (P) and DNA represents the DNA fragment with a defined number of binding sites. This can be rearranged to determine a dissociation constant (K_d):

$$K_d = \frac{[P]^n\,[DNA]}{[Pn\cdot DNA]}$$

Taking the log of this gives:

$$\log K_d = n\log[P] + \log\left(\frac{[DNA]}{[Pn\cdot DNA]}\right)$$

This can then be rearranged to give:

$$\log\left(\frac{[DNA]}{[Pn\cdot DNA]}\right) = -n\log[P] + \log K_d$$

Graphing the log of the concentrations of free DNA divided by the complexed DNA, versus the log of the protein concentration is a Hill plot and the slope (–n) is also called the Hill coefficient (Mathews and Van Holde, 1990, pp. 222-225).

Bending and competition assays. For bending assays PCR-amplified DNA fragments (approximately 50 000 cpm) were digested with a series of restriction enzymes for 8 hours. Digests were extracted with phenol and chloroform, ethanol precipitated and resuspended in 20 μL of Milli-Q® water. 2 μL of the digested DNA was used in binding reactions, run on a 8% polyacrylamide gel, and put on a Phosphor Screen (Molecular Dynamics) overnight. The relative mobility of the *Eco*RV digested and retarded band was calculated in comparison to the *Bam*HI digested and retarded band. The bend angle was then calculated by using the formula $\alpha=2[\cos^{-1}(\text{relative mobility})]$ (Kim *et al*., 1989).

The binding sites cloned in pRF911 (*sbmA*) and pRF940 (*sbmABC*) were amplified by PCR using primers RFE16 and RFE17 and used as competitor DNA in the competition assays. This non-radioactive DNA was run on a 1.5% agarose gel, excised, and quantified by A_{260}. TraM and radio-labelled DNA was bound at 37°C for 15 minutes as in the retardation reactions. 1μL of competitor DNA was added to the reactions and incubated at 37°C for another 15 minutes. Reactions were run on a 5% acrylamide gel, which was subsequently dried and placed on a Phosphor Screen (Molecular Dynamics) overnight.

Hydroxyl radical footprinting. Approximately 0.5 μg of plasmid DNA (pRF911, pRF920) was digested with *Pvu*II and run on a 1.8% agarose gel. The 156 bp and 183 bp (respectively) fragments were excised from the gel and purified. The DNA was eluted in 50 μL of dH_2O and 10 μL was used in a sequencing reaction using either RFE9 (GCTGCCCGGGAGGCCTTC) or RFE10 (GCTGGATATCTTTAAACTCGAG), Sequenase (USB) and ^{33}P-labelled ddNTP (Pharmacia). Templates used for the hydroxyl-radical reactions were created by end-labelling 125 pmol of RFE9 or RFE10 using 5 μL (50Ci) of $[\gamma^{32}P]$ATP (ICN) and Polynucleotide Kinase (Roche) for 1 hour. Primers were purified using Quick Spin Oligo Columns (Roche) and eluted with approximately 50 μL Milli-Q® water. PCR reactions using the end-labelled primer and 125 pmol of either RFE9 or RFE10 were then performed. The reactions were dried down to 15 μL using a concentrator (Savant) and loaded onto a 1.8% agarose gel. The 158 bp and 185 bp (respectively) fragments were excised and purified. Radioactivity was quantitated using a Scintillation counter LS3801 (Beckman) and approximately 5000 cpm was used in each reaction. Purified TraM was allowed to bind to the DNA in a 15 µL volume in TE buffer for 15 minutes at 37°C. On the side of the tube 2 µL of each of 10mM sodium ascorbate, 0.2 mM iron ammonium sulfate, 0.4 mM EDTA, and 0.3 % H_2O_2 (v/v) were mixed and allowed to drop into the binding reaction. After 2 minutes 12 µL of Stop Solution (32 µL of 0.2 M EDTA, 10 µL of 0.1 M thiourea, and 1 µL of 0.5 mg/mL tRNA) was added with 3 µL of 100 % glycerol to stop the cleavage reaction. Reactions were loaded onto a 5 % TB polyacrylamide gel and run at 28 mA at 4°C until the bromophenol blue dye reached the bottom of the gel. The gel was wrapped in Saran-Wrap and exposed to X-OMAT AR film (Kodak) for 4 hours. After development of the film, bands which correlated with bound and unbound DNA were excised. The DNA was eluted overnight in 400 µL of 0.5 M ammonium acetate, 1mM EDTA at 37°C. The tubes were then centrifuged at 14 000 rpm in a bench top centrifuge. 1 mL of 95 % ethanol was added to the aqueous layer to precipitate the DNA. The DNA was dissolved in 5 µL of Sequencing Stop Solution (USB) and loaded onto an 8 % polyacrylamide gel containing 8 M urea. As a marker, 1 µL of the 10 µL G, A, T, and C sequencing reactions were loaded onto the gel. The gel was run at 40 W until the xylene cyanol marker was approximately 5 cm from the bottom, dried and was exposed to a Molecular Dynamics Phosphor Screen for up to 5 days.

Chapter 3

Analysis of the allele specific nature of *oriT*

A version of this was reprinted as Fekete and Frost, 2000. J. Bacteriol. **182**:4022-4027

Introduction

Conjugation is the horizontal transfer of DNA from donor to recipient bacteria via plasmid-derived transfer (*tra*) proteins and other host-encoded factors. F, R1, and R100-1 (a derepressed mutant of R100; Anthony *et al.*, 1999) are closely related members of the IncF group of self-transmissible plasmids (Frost *et al.*, 1994) which exhibit plasmid specificity (Willetts and Maule, 1986). The plasmids are differentiated at the level of transcriptional control of the major *tra* operon as well as by properties associated with the conjugative pilus including antigenicity, phage sensitivity, entry exclusion and mating pair stabilization (Anthony *et al.*, 1999). The *oriT* region contains the site (*nic*) for strand and sequence-specific cleavage by the relaxase (TraI), as well as binding sites for the plasmid-specific TraM and TraY proteins, and the host-encoded IHF (Integration Host Factor; Fig. 1.2). Removal of any of these binding sites decreases the mobilization efficiency of *oriT*-containing chimeric plasmids (Abo and Ohtsubo, 1995; Fu *et al.*, 1991). Plasmid specificity of the transfer gene products for their cognate *oriT* has been demonstrated for TraI and TraM (Everett and Willetts, 1980; Willetts and Maule, 1986).

IHF binds two sites in both the F (Tsai *et al.*, 1990) and R100 *oriT* region (Dempsey and Fee, 1990; Inamoto *et al.*, 1990). Both intrinsic bends and bends induced by IHF (Tsai *et al.*, 1990) are thought to fulfill the 3-D structural requirements at *oriT* necessary for cleavage at *nic* and interaction with the transferosome prior to transfer. The position of *nic* has been established for F (Sherman and Matson, 1994; Thompson *et al.*, 1989) and R100-1 (Inamoto *et al.*, 1991), which are equivalent except for a two base pair difference in the sequence immediately adjacent to *nic* (Frost *et al.*, 1994). The coupling protein which links the relaxosome to the transferosome is thought to be TraD, an inner membrane protein (Achtman *et al.*, 1979; Sastre *et al.*, 1998).

TraY, encoded by the first gene in the *traY-I* operon, binds at two sites in F *oriT* (*sbyA* and *sbyC*; Lahue and Matson, 1990; Luo *et al.*, 1994), and one site in *oriT* of R100 (*sbyA*; Inamoto and Ohtsubo, 1990). The relaxase, TraI, cleaves a single-strand of DNA in *oriT* at a site now called *nic* and covalently binds to the 5' end (Byrd and Matson, 1997; Inamoto *et al.*, 1991; Lanka and Wilkins, 1995; Matson and Morton, 1991). In addition to relaxase activity, F TraI also contains an ATP-dependent helicase activity in the large carboxyl-terminal domain of the molecule (Dash *et al.*, 1992). TraI has been localized to the cytoplasm (Achtman *et al.*, 1979) but upon overexpression in the presence of TraD has been shown to be associated with the inner membrane (Dash *et al.*, 1992). In the F plasmid, IHF and TraY are required for the nicking reaction *in vitro* (Nelson *et al.*, 1995), and assembly of the resulting "relaxosome" occurs in a specific order with TraI binding after IHF and TraY (Howard *et al.*, 1995). Similar characteristics have been shown for the closely related plasmid R100 (Abo and Ohtsubo, 1995; Fukuda and Ohtsubo, 1997).

TraM is a cytoplasmic protein of 14.5 kDa which forms tetramers in solution (Frost *et al.*, 1997; Verdino *et al.*, 1999). It binds to three sites in F *oriT* (*sbmA*,*B*,*C*; DiLaurenzio *et al.*, 1992) and four sites in *oriT* of R100 (*sbmA-D*; Abo *et al.*, 1991). In F, one of these sites, *sbmC*, is associated with transfer while the other two, *sbmA* and *sbmB*, are involved in the autoregulation of *traM* transcription (Penfold *et al.*, 1996). Removal of *sbmA* and *sbmB* (Fig. 1.2), decreases mating efficiency 100-fold, while further deletion of *sbmC* results in an additional 100-fold decrease in mobilization efficiency of a plasmid containing a cloned version of *oriT* (Fu *et al.*, 1991). TraM from the F-like plasmids R100-1 (Abo and Ohtsubo, 1993), and R1 (Schwab *et al.*, 1993) also autoregulates transcription. The amino-terminal region of TraM is responsible for DNA binding (Frost *et al.*, 1997; Schwab *et al.*, 1993) in a plasmid-specific manner (Kupelwieser *et al.*, 1998). The F and R100-1 TraM proteins are 127 amino acids long and are 88.9% identical and 95.3% similar with 11 of the 14 differences occurring in the first 37 amino acids of the proteins. TraM has also been shown to be associated with the inner membrane (Achtman *et al.*, 1979; DiLaurenzio *et al.*, 1992) possibly via the inner membrane protein TraD (Disque-Kochem and Dreiseikelmann, 1997).

Recent evidence demonstrating that TraM (but not TraY) is required in conjunction with IHF and TraI for cleavage at *nic* in R1 (Kupelwieser *et al.*, 1998) contradicts existing *in vitro* data for F (Howard *et al.*, 1995). This data demonstrates that F TraM is not essential for cleavage at *nic* suggesting that, despite the high homology between F and R1 plasmids, the control of the cleavage reaction can vary considerably between these two plasmids.

Previous work has shown that TraM, TraY, and TraI from F, R100-1, and R1 plasmids showed plasmid specificity for their homologous *oriT* regions, with TraM and TraY thought to have more specificity than TraI based on sequence variation and number of alleles (Frost *et al.*, 1994). Due to the clear differences between the F and R100-1 mating pair formation systems (Anthony *et al*., 1999) and the plasmid specificity exhibited by the transfer proteins that bind *oriT*, chimeric plasmids that are hybrids of the F and R100-1 *oriT* regions were constructed. These results indicate that there are complex interactions between the proteins that bind at *oriT* that define the efficiency of cleavage at *nic* and that ensure correct interaction with the transferosome prior to transfer.

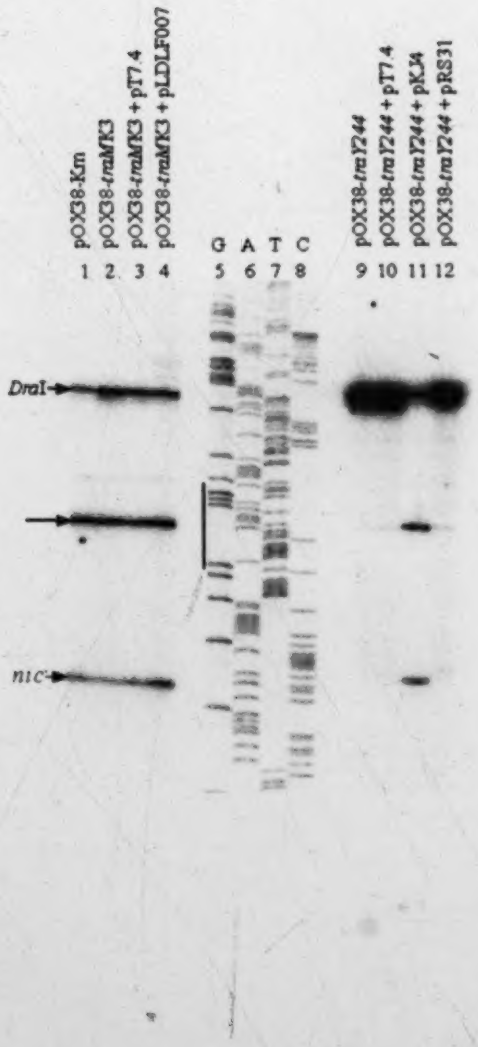
Contrary to previous data using multicopy plasmids, neither TraM nor TraI were found to be associated with the membrane. The presence of amber mutations in *traD* or *traG* did not affect the localization pattern of TraM, suggesting that, in the absence of recipient cells, the relaxosome might not be associated with the transferosome and that the signal that triggers conjugative DNA transfer might involve localization of the relaxosome to the base of the pilus prior to transfer.

Results

In order to assess the role of TraM in promoting cleavage at *nic*, a nicking assay was performed in *E. coli* XK1200 using pOX38-Km (which contains the entire wild-type F transfer region) and its derivatives pOX38-*traM*K3 (*traM*) and pOX38-*traY244* (*traY*) (Figure 1.2). In agreement with Howard *et al.* (1995), removal of TraY by mutation in plasmid pOX38-*traY244* abolished cleavage at *nic* (Figure 3.1, lane 9). This could be due to either a requirement for TraY as part of the relaxosome or to the positive regulatory effect of TraY on the P_Y promoter which transcribes the *traI* gene as part of the 33 kb *tra* transcript. Supplying TraY in *trans* (pKJ4) restored cleavage at *nic* (Figure 3.1, lane 11), while the addition of the *traI* gene in *trans* (pRS31) did not (Figure 3.1, lane 12), consistent with previous data that showed that TraI requires TraY for its expression and relaxase activity.

The level of cleavage at *nic* in the *traM* mutant, pOX38-*traM*K3, was equivalent to the wild type plasmid pOX38-Km (Figure 3.1, lanes 1 and 2). A slight increase in the level of cleavage was seen upon the addition of extra TraM in *trans* (pLDLF007; *traM* transcribed from its own promoters) compared to the vector control (pT7.4) (Figure 3.1, lanes 3 and 4). This suggested that the level of nicking could be further maximized by increasing the intracellular amount of TraM. A band was routinely found between *nic* and the *Dra*I site (Figure 3.1, middle arrow) which was within the AT-rich region containing IHFA, the first IHF binding site (Tsai *et al.*, 1990). The intensity of this band

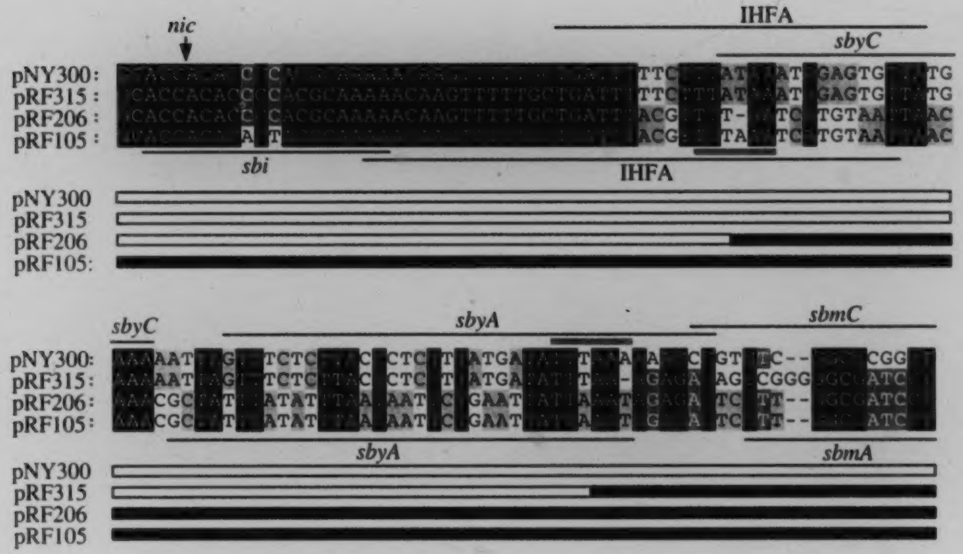
Figure 3.1. Nicking reactions of pOX38-Km and its derivatives. Plasmids present in each experiment are listed above each lane. The *Dra*I site and the cleavage site (*nic*) are indicated with arrows. The sequencing ladder is used to identify the *nic* and *Dra*I sites and was performed using the same primer as in the nicking reactions. A non-specific band is identified with an arrow between *nic* and the *Dra*I site. The IHF binding site (IHFA) is designated by a vertical line next to the G lane in the sequencing reaction.


pOX38-Km
pOX38-*traMK3*
pOX38-*traMK3* + pT7.4
pOX38-*traMK3* + pLDLF007
G A T C
pOX38-*traY244*
pOX38-*traY244* + pT7.4
pOX38-*traY244* + pKJ4
pOX38-*traY244* + pRS31
1 2 3 4 5 6 7 8 9 10 11 12
DraI
nic

reflected the level of cleavage at *nic* and was greatly reduced when TraY was absent (Figure 3.1, lane 9) suggesting that termination at this site is dependent on relaxosome formation. Since IHF has been found to bend DNA at 140° (Thompson and Landy, 1988) this site might be susceptible to breakage during preparation of the sample. The increase in the intensity of this band when cleavage occurs at *nic* (Figure 3.1, lane 11) suggests that IHF might cause increased strain on the DNA at this point when the relaxosome has formed. In the absence of cleavage at *nic*, the band located at the *Dra*I site was intensified as expected (Figure 3.1, lane 9) and was approximately equivalent to the sum of the intensities of bands at *nic*, IHFA, and *Dra*I in other samples.

Since the sequences between the F and R100-1 TraM and TraY binding sites within the *oriT* region are highly conserved, chimeric plasmids were constructed from PCR products or restriction fragments of *traM*, the TraM and TraY binding sites and the *nic* site as described in Materials and Methods (Figure 3.2). Fortuitously situated *Dra*I sites between *sbyA* and *sbmC* in F, and between *nic* and *sbyA* in R100-1 allowed construction of hybrid plasmids in which *nic* and *sbyA* as well as the TraM region were shuffled. pNY300 (Frost *et al*., 1989) contains the F *oriT* (*nic*, IHFA, *sbyA*, and *sbmABC*) and the F *traM* gene. pRF105 contains the R100-1 *oriT* (*nic*, IHFA, *sbyA*, and *sbmABCD*) and the R100-1 *traM* gene. pRF315 was constructed by linking the F *nic* and TraY binding site (*sbyA*) to the R100-1 TraM binding sites (*sbmABCD*) and the *traM* gene. It had one base pair missing from the *Dra*I site and two additional base pairs near the beginning of sbmC which did not affect its ability to be mobilized by pOX38-Km or pOX38*traM*K3 when compared to the mobilization frequency of pNY300 (see below). pRF206 was

Figure 3.2. Sequences of the *oriT* regions of the chimeric plasmids. Sequences are aligned at *nic* according to Frost *et al.* (1994). Binding sites for F proteins and IHF are shown above the pNY300 sequence while the equivalent binding sites for R100-1 are represented below the pRF105 sequence. *nic* is identified by arrows above and below the sequences. Sequences were compared by PILEUP in GCG and 100% homology is represented by black boxes, 75% with gray boxes with white lettering, 50% with gray or white boxes with black lettering. The *Dra*I sites used for the cloning of pRF315 and pRF206 are shown as dark gray lines above and below the pNY300 and pRF105 sequences, respectively. Below the sequences is a diagram of the F (clear boxes) and R100-1 (black boxes) sequences for each chimeric *oriT* region.


nic
IHFA
sbyC
pNY300:
pRF315:
pRF206:
pRF105:
sbi
IHFA
pNY300
pRF315
pRF206
pRF105:
sbyC
sbyA
sbmC
pNY300:
pRF315:
pRF206:
pRF105:
sbyA
sbmA
pNY300
pRF315
pRF206
pRF105

constructed by linking the F *nic* to the R100-1 TraY and TraM binding sites (*sbyA* and *sbmABCD*) and the *traM* gene. One base pair was removed from the *Dra*I site during construction which did not affect its mobilization frequency by R100-1 when compared to pRF105 (see below).

The chimeric plasmids (Figure 3.2) were tested for their ability to be mobilized by R100-1, F (pOX38-Km), and an F *traM* mutant, pOX38-*traM*K3 (Table 3.1A). They were also tested for their ability to complement pOX38-*traM*K3 as well as their effect on the transfer of the pOX38-Km plasmid (Table 3.1B). For mobilization assays, pOX38-Km or pOX38-*traM*K3 in *E. coli* DH5α were used as donor cells and ED24 as recipient cells. Mobilization assays in the presence of R100-1 used *E. coli* JE2571-1/R100-1 as donor cells and CS2198 (Km^r) as recipient cells.

pNY300 was mobilized in the presence of pOX38-Km and pOX38-*traM*K3 (since pNY300 supplies F TraM), but not in the presence of R100-1 (Table 3.1A). Similarly, pRF105 was mobilized by R100-1 but not by pOX38-Km or pOX38-*traM*K3, a finding which is consistent with the previously determined plasmid specificity of TraM for its cognate *oriT* region (Willetts and Maule, 1986).

pRF315 was mobilized efficiently only in the presence of pOX38-Km and pOX38-*traM*K3 (51 and 6.7 transconjugants per 100 donors, respectively; Table 3.1A). Since TraM is required for transfer, the R100-1 TraM supplied by pRF315 was able to bind to the R100-1 *sbmABCD* sites on pRF315 and interact with the F *tra* proteins supplied by

Table 3.1A. Mobilization of chimeric plasmids

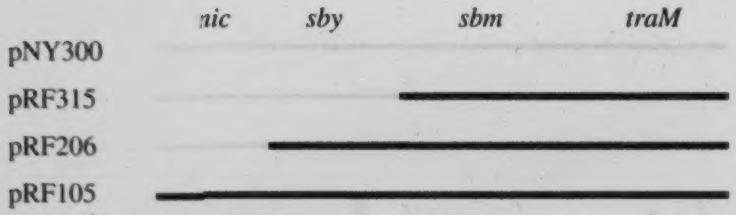
	Number of transconjugants/100 donors with chimeric plasmid:				
	pNY300	pRF315	pRF206	pRF105	pUC18
pOX38-Km[a]	110	51	71	0	0
pOX38-*traM*K3	7.0	6.7	5.2	0.008	0
R100-1	0.004	0.009	6.4	28	0

[a] Column contains plasmids that supply transfer functions

Table 3.1B. Effect of chimeric plasmids on transfer of pOX38-Km, pOX38-*traM*K3

	Number of transconjugants/100 donors with chimeric plasmid:				
	pNY300	pRF315	pRF206	pRF105	pUC18
pOX38-Km[a]	97	83	36	35	133
pOX38-*traM*K3	3.7	0.007	0	0	0

[a] Column contains plasmids that supply transfer functions



Chimeric plasmids used in tables as described in Figure 3.2. Grey-F plasmid, Black-R100-1 the pOX38 plasmids. The lack of mobilization of pRF315 by R100-1 (less than 0.01 transconjugants per 100 donors) suggests that the R100-1 relaxosome and/or transfer apparatus was not able to interact efficiently with the F *nic* and *sbyA* sequences found on pRF315.

pRF206 was mobilized in the presence of pOX38-Km and R100-1 (71 and 6.4 transconjugants per 100 donors, respectively) suggesting that both F and R100-1 are able to transfer this construct at approximately the same level that they mobilized pNY300 and pRF105, respectively. Interestingly, the presence of F TraM supplied by pOX38-Km increased the levels of mobilization approximately 10-fold suggesting that heteromultimers of TraM interact more efficiently with the F transferosome than R100-1 TraM alone. The decrease in mobilization efficiency for pRF206 compared to pRF105 in the presence of R100-1 suggests that the presence of the F *nic* site in pRF206 did have a minor negative affect on mobilization.

Mating efficiency assays showed that pNY300 (supplying F TraM), but not pRF315, pRF206 or pRF105 (supplying R100-1 TraM), was able to complement the *traM* mutation in pOX38-*traM*K3 (Table 3.1B). This suggests that TraM must bind in *cis* to *nic* for transfer to occur. pOX38-Km transferred at normal levels in the presence of all four chimeric plasmids (Table 3.1B), suggesting that the presence of R100-1 TraM did not exert a dominant negative effect on F TraM function.

The phenotypes of cleavage and transfer have been used to define whether the relaxosome is stable and whether it is able to interact with the transfer machinery (transferosome) to effect DNA transfer (Fu *et al*., 1991). To differentiate between these possibilities, nicking assays were performed on the four chimeric plasmids in both F and R100-1 backgrounds (Figure 3.3, Table 3.2). The unidirectional amplification procedure involved binding a primer to a sequence within pUC18 to generate a single-stranded product which terminated at either *nic* or at a restriction enzyme site downstream from *nic*. Since the primer annealed to the pUC18 vector, no product derived from the F or R100-1 co-resident plasmids was obtained. *Dra*I was used to terminate the products for plasmids pNY300 and pRF315 (91 bases from *nic*), while *Hin*fI was used for pRF206 and pRF105 (98 bases from *nic*). Band intensities were quantitated using Image QuaNT and the ratio of the intensity of the band at *nic* to the sum of the bands in each sample is given as a percentage (Table 3.2). Values in Table 3.2 that were less than or equal to 0.1% were assumed to be insignificant and most likely were due to inherent errors during quantification of band intensities. An extra band (hatched arrow, Figure 3.3) immediately above *nic* and not associated with the band at IHFA, was routinely seen in all samples and was considered to be an artifact of sample preparation using the high copy number vector pUC18. The use of pUC18 as the cloning vector allowed very low levels of nicking to be visualized by this assay, but also, seemed to increase the likelihood of DNA damage during preparation.

In agreement with mating efficiency results, pNY300 was cleaved at *nic* in the presence of pOX38-Km and pOX38-*traM*K3 but not R100-1 (Table 3.2), while pRF105 was

Figure 3.3. Examples of nicking reactions using the chimeric plasmids from Figure 3.2 Lanes 1-4, 9-12, 17-20 and 23-26 are the sequencing reaction for each plasmid using the universal primer (G,A,T,C respectively). Lanes 5-8 are the nicking reactions for pNY300 and lanes 13-16 for pRF315 in DH5α alone and with pOX38-Km, pOX38-*traM*K3 and R100-1, respectively. Lanes 21 and 22 are the nicking reactions for pRF206 with pOX38-Km and R100-1, respectively. Lanes 27-29 are the nicking reactions of pRF105 with pOX38-Km, pOX38-*traM*K3, and R100-1, respectively. The *nic* and restriction enzyme sites are identified with arrows. An example of a non-specific band is identified with a hatched arrow.

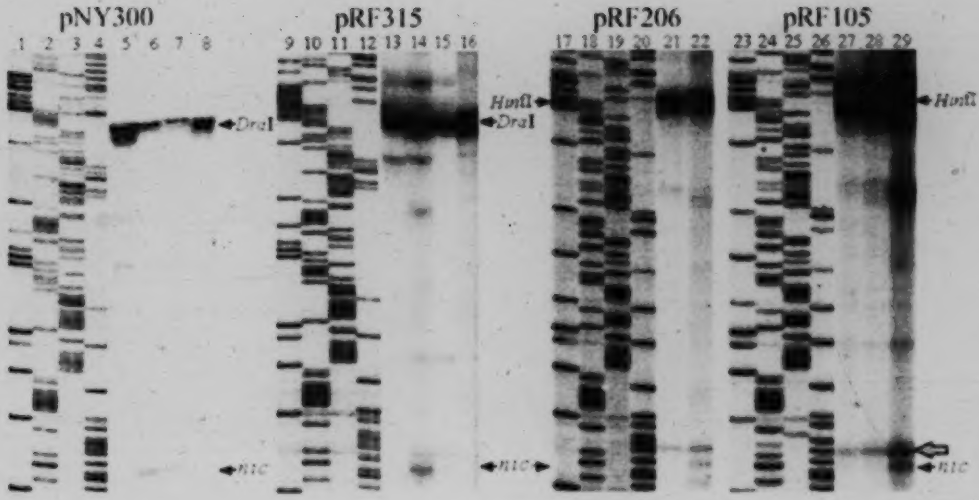

pNY300
1 2 3 4 5 6 7 8
DraI
nic
pRF315
9 10 11 12 13 14 15 16
HinfI
DraI
nic
pRF206
17 18 19 20 21 22
pRF105
23 24 25 26 27 28 29
HinfI
nic

Table 3.2. Percentage of chimeric plasmids cleaved at *nic*

	Percent of cleavage at *nic* in each chimeric plasmid:			
	pNY300	pRF315	pRF206	pRF105
pOX38-Km[a]	17	12	0.2	0.1
pOX38-*traM*K3	12	4	0.3	0.1
R100-1	0	0	3	2
DH5α[b]	0	0	0	0

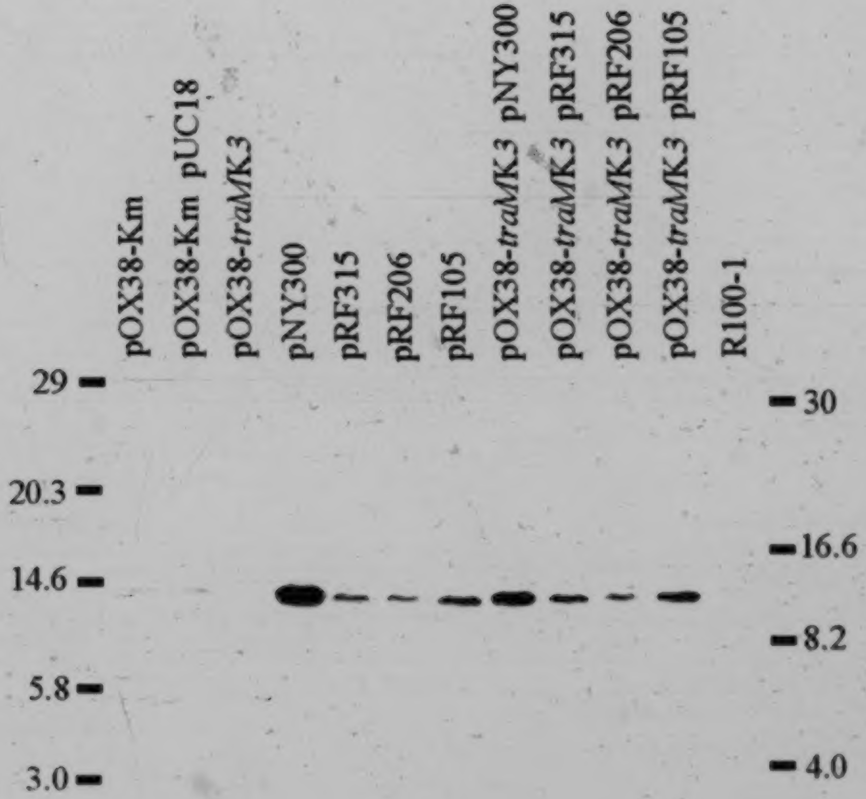
[a] Column contains plasmids that supply transfer functions

[b] Cells contain no transfer proteins

efficiently cleaved only in the presence of R100-1. pRF315 was cleaved by pOX38-Km and to a lesser extent by pOX38-*traM*K3 in agreement with the mobilization results for this plasmid (Table 3.1A). Unexpectedly, pRF206 was efficiently cleaved in the presence of R100-1 but cleavage was almost undetectable in the presence of pOX38-Km or pOX38-*traM*K3. Cleaved species of pRF206 are presumed to be present since relaxation at *nic* is required for transfer, and since mobilization of pRF206 was roughly comparable to pNY300 and pRF315 in an F (pOX38) background.

R100-1 and F TraM share 95.3% similarity and 89% identity at the amino acid level, and eleven of the 14 differences occur in the first 39 amino acids. Based on this high degree of identity it was hoped that anti-F TraM antisera would be able to detect R100-1 TraM. In order to confirm that TraM (F or R100-1) was being expressed in all of the chimeric plasmids, cell lysates were run on a 15% SDS polyacrylamide gel and Western analysis was performed after transfer (Figure 3.4). TraM can be seen in pOX38-Km with and without pUC18 showing that expression of TraM is not affected by the presence of this high copy vector. As expected TraM was not present in the *traM* mutant, pOX38-*traM*K3, but was present in pNY300 (the clone containing the F *traM* gene). Interestingly, R100-1 TraM was detected by the F TraM antisera in pRF315, pRF206, and pRF105. TraM was also detected from all of the chimeric plasmids when the F *traM* mutant, pOX38-*traM*K3, was present in the same cell. This confirms that TraM is present in the cell during the nicking and mobilization experiments using the chimeric plasmids.

Figure 3.4. Immunoblot using Anti-F TraM antiserum to detect the presence of TraM in cells is shown. The plasmids present in the cells are designated at the top of each lane. Approximately 0.1O.D.$_{600}$ of each cell type was run in each lane. Bars on each side of the figure show the approximate position and size in kDa of the size standards run in the gel. Upon longer exposure protein was detected in the R100-1 lane.


pOX38-Km
pOX38-Km pUC18
pOX38-*traMK3*
pNY300
pRF315
pRF206
pRF105
pOX38-*traMK3* pNY300
pOX38-*traMK3* pRF315
pOX38-*traMK3* pRF206
pOX38-*traMK3* pRF105
R100-1
29
20.3
14.6
5.8
3.0
30
16.6
8.2
4.0

The above experiments demonstrate that F TraM is not required for cleavage at *nic*, but is required for subsequent steps in the mobilization of DNA. It has been postulated that TraM anchors the relaxosome to the inner membrane possibly via interaction with TraD. If TraM is peripherally associated with the membrane through TraD then mutations in *traD* should affect this localization. The presence of TraM, TraI and TraD were assayed in cell fractions consisting of soluble (cytoplasmic plus periplasmic) and insoluble (membrane) fractions by immunoblot. Crude membranes were further washed to assay for peripheral association of the proteins with the membrane.

Cytoplasmic contamination of membrane fractions was measured by assaying for glucose-6-phosphate dehydrogenase activity. Lowry assays (Ausubel *et al*., 1989) were used to determine protein concentrations for SDS polyacrylamide gel analysis as well as for the determination of glucose-6-phosphate dehydrogenase specific activity (Table 3.3).

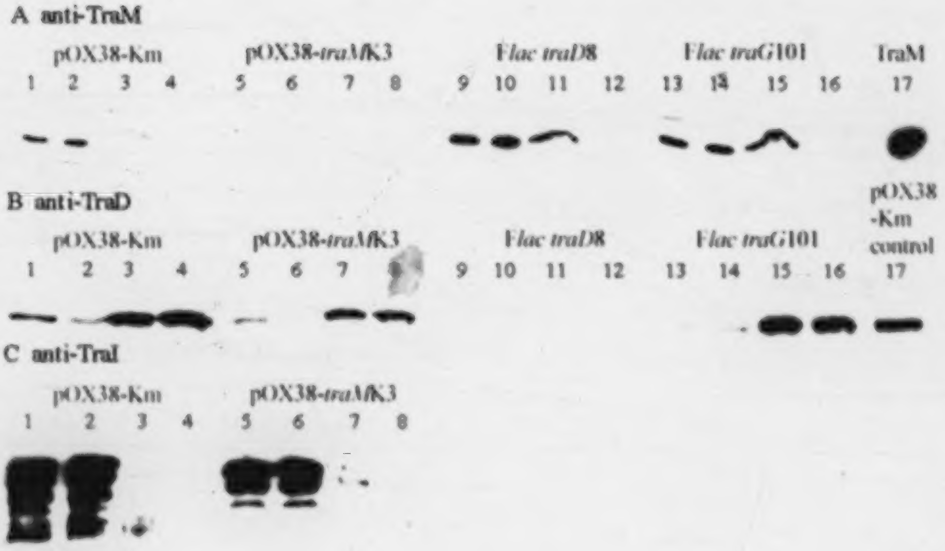
Detection of TraM, TraD, and TraI by immunoblot was performed using cells containing pOX38-Km, pOX38-*traM*K3, and F*lac traD8* which carries an amber mutation in the 103 codon of *traD* (Anthony, K., *et al*., 1999; Willetts and Achtman, 1972). Another mutant F plasmid, F*lac traG101*, which contains a frameshift mutation in *traG* and is transfer-deficient (Achtman *et al*., 1972) was also tested. TraG is an inner-membrane protein containing a large periplasmic domain (Firth and Skurray, 1992) and is the only other major inner membrane protein in F suggested to be involved in linking the relaxosome to the transferosome (Anthony *et al*., 1999). Immunoblot analysis of TraM in each of the fractions from each strain was performed as shown in Figure 3.5a. As expected, TraM

Table 3.3. Glucose-6-phosphate dehyrogenase activity

Strain/Plasmid	Lysate	Cytoplasm/ periplasm	Unwashed membrane	Washed membrane
DH5α / pOX38-Km	369	749	18	118
DH5α / pOX38-*traM*K3	331	499	0	36
ED2149 / F*lac traD*8	462	858	0	0
ED2149 / F*lac traG*101	455	919	53	0

[a] Data given as mol/total mg protein/minute

Figure 3.5. Protein localization by immunoblot analysis. Anti-TraM (A), -TraD (B), and -TraI (C) immunoblots for pOX38-Km (lanes 1-4), pOX38-*traM*K3 (lanes 5-8), F*lac traD8* (lanes 9-12), and F*lac traG101* (lanes 13-16) as described in Materials and Methods. Lanes 1, 5, 9 and 13, whole cell lysates; 2, 6, 10 and 14, cytoplasmic and periplasmic fractions; lanes 3, 7, 11 and 15, unwashed membrane fractions; lanes 4, 8, 12 and 16, washed membrane fractions. In lane 17 of A, 22 ng of purified TraM was run as a positive control. In lane 17 of B, 0.4 O.D.$_{600}$ of pOX38-Km cells was run as a positive control (see Materials and Methods).


A anti-TraM
pOX38-Km
pOX38-*traM*K3
F*lac traD*8
F*lac traG*101
TraM
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
B anti-TraD
pOX38-Km
pOX38-*traM*K3
F*lac traD*8
F*lac traG*101
pOX38 -Km control
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
C anti-TraI
pOX38-Km
pOX38-*traM*K3
1 2 3 4 5 6 7 8

was principally found in the cytoplasmic/periplasmic fraction with a small amount in the unwashed membrane fraction for pOX38-Km. The association of TraM with the membrane fractions of F*lac traD8* and F*lac traG101* suggests its weak association with the membrane may involve other proteins. Since it is easily washed out of the membrane fraction (wash with 10% sucrose and 5 mM EDTA), it is most likely not a true peripheral membrane protein. This suggests that the relaxosome is not associated with TraD via TraM during vegetative growth, or is not detectable using these assays. As previously reported (Panicker and Minkley, 1992), TraD was primarily found in the membrane fraction with a small amount found in the cytoplasm (Figure 3.5b) while TraI was found in the soluble cytoplasmic fraction (Figure 3.5c).

Discussion

TraI, TraY and IHF are required for cleavage at *nic in vitro* in the F plasmid (Howard *et al.*, 1995) while TraM, rather than TraY, is required for cleavage at *nic in vivo* in the F-like plasmid R1 (Kupelwieser *et al.*, 1998). In accordance with the *in vitro* data for F, we have found that F TraY and TraI are required for cleavage at *nic in vivo* while F TraM is not essential. Since TraY is required for TraI expression by virtue of its regulatory role at the Py promoter which transcribes the *tra* operon (*traY-I*; Frost and Manchak, 1998), the *traY244* mutation could affect nicking indirectly by blocking *tra* operon expression. Since supplying pRS31, which is a source of TraI but not TraY, did not restore cleavage at *nic*, there appeared to be a requirement for expression of the complete *tra* operon. Complementation of the *traY244* mutation with pKJ4 (TraY) restored both pilus expression (Anthony, 1998) and cleavage, suggesting that TraY and TraI are required for cleavage (no other protein expressed by the *tra* operon has been implicated). Everett and Willetts (1980) have shown that TraM is not required for cleavage at *nic* using an *in vivo* lambda nicking assay. Similarly, Achtman *et al.* (1972) showed that a mutation in *traM* (JCFL102) affected transfer ability but not phage sensitivity, an accurate method for detecting pilus formation and *tra* operon expression. In Kingsman and Willetts (1978), the *traM102* mutation was shown to affect the initation of DNA synthesis in the donor after mating pair formation had occurred. Since R1 TraM has both a regulatory role in the expression of pili (Poltzleitner *et al.*, 1997) and in the level of nicking (Kupelwieser *et al.*, 1998) there appears to be interesting differences between these two apparently homologous systems.

The requirement of F TraM for efficient transfer but not for efficient cleavage suggests that TraM acts at a stage after relaxation of *nic* within the F *oriT* to promote transfer as suggested by the results of Kingsman and Willetts (1978). Interestingly, supplying TraM in *trans* from a multicopy plasmid increased the amount of cleavage at *nic* suggesting that the equilibrium between nicked and un-nicked DNA was shifted towards the relaxed species.

The organization of the *oriT* region of the R100-1 plasmid closely resembles that of the F plasmid with plasmid specificity being defined at the level of TraI, -M and -Y binding at their cognate sites within *oriT*. Once binding to the DNA has taken place, further specificity could be provided by protein-protein interactions between these proteins within the relaxosome as well as with other proteins involved in the transfer process. Thus, the level of relaxation at *nic* could reflect the ability of TraY to bind the *oriT* region independently of TraI (for example: if TraY alters the conformation of the DNA near *nic* thereby affecting TraI function), or reflect the presence of direct interactions between TraY and TraI. Similarly, the interaction of TraM with these proteins, as well as interactions between the relaxosome and transfer apparatus, could also define plasmid specificity. Chimeric plasmids, containing portions of the F and R100-1 plasmids, were constructed to identify possible interactions between these proteins at a level other than simple DNA recognition. This system allowed the definition of the factors required for efficient cleavage as opposed to efficient transfer by the respective transfer systems of F and R100-1.

The R100-1 TraM protein of pRF315, pRF206, and pRF105 was not able to complement the *traM* mutation in the F plasmid derivative pOX38-*traM*K3. This was not due to decreased cleavage at *nic* since TraM is not required for this step in F transfer, but may be due to the inability of R100-1 TraM to bind the F TraM binding sites. Since purified F TraM has a low affinity for R100-1 TraM binding sites as measured by EMSA (appendix), R100-1 TraM might also have a correspondingly low affinity for F TraM binding sites. If this assumption is correct, TraM must be bound to sites in *cis* to *nic* for the relaxosome complex to be directed to the transfer apparatus. This is further supported by the evidence that pOX38-*traM*K3 can efficiently mobilize pRF315, where R100-1 TraM is bound to its cognate sites in *cis* to F *nic*.

Using *in vitro* affinity assays, TraM has been shown to interact with TraD (Disque-Kochem and Dreiseikelmann, 1997), which is thought to couple the relaxosome to the transferosome (Sastre *et al*., 1998). Thus, TraM could act as the link between TraD and the transferosome and the plasmid that has been readied for transfer. Since only the N-terminal regions of F and R100-1 TraM vary to any significant degree, the domain responsible for interacting with the transfer apparatus probably resides within the homologous C-terminal domains. TraD does not show plasmid specificity among the F-like plasmids (Willetts and Maule, 1986), however, deletions at its C-terminus affect its ability to act as a coupling protein (Sastre *et al*., 1998). It would be interesting to determine whether the C-termini of TraM and TraD do, in fact, interact as predicted from these results.

Since the level of transfer of pOX38-Km was unaffected by the presence of the chimeric plasmids supplying R100-1 TraM, there appeared to be no dominant negative effect resulting from having both types of TraM within the same cell. Thus, either mixed oligomers are fully functional or F TraM is preferentially selected to bind to F *oriT*, supplying further evidence for a mechanism that guarantees plasmid specificity during transfer.

The F relaxase, TraI, appears to cleave the F *nic* site of pRF206, containing a R100-1 TraY binding site, since mobilization remains efficient in the presence of pOX38-Km. However, cleavage is only detectable, both *in vitro* (Nelson *et al.*, 1995) and efficiently *in vivo* when F TraY is bound in *cis* to *nic* (as in pRF315). This imbalance between cleavage and mobilization suggests that the few copies of pRF206 that have been cleaved by F TraI are preferentially selected by the transferosome for transfer.

Interestingly, no cleavage or transfer by the F (pOX38-Km) transfer system is observed for pRF105 which contains a two base pair difference in sequence near *nic*. This suggests that the F TraI relaxase cannot cleave at R100-1 *nic* and can discriminate between F and R100-1 *nic* on the basis of this sequence difference.

In the presence of R100-1, which supplies R100-1 TraI, both pRF105 (R100-1 *nic*) and pRF206 (F *nic*), were cleaved and mobilized at comparable levels. Substitution of the R100-1 TraY binding site with F sbyA, as in pRF315, resulted in no cleavage or mobilization by R100-1. Thus R100-1 TraI cleaved F *nic* if R100-1 TraY was bound in *cis* to its cognate binding site (pRF206) suggesting that TraY provides another level of specificity in the R100-1 system.

A specific function for F TraM has not yet been defined. TraM is essential for transfer (Achtman *et al.*, 1971; Penfold *et al.*, 1996) and its ability to bind DNA near *nic* and interact with TraD (Disque-Kochem and Dreiseikelmann, 1997) suggests that TraM may anchor the DNA to the membrane. TraM has also been proposed to promote relaxosome formation via formation of a nucleosome-like structure at *oriT* which adjusts the superhelical density and promotes cleavage and unwinding in preparation for transfer (Kupelwieser *et al.*, 1998; Penfold *et al.*, 1996). The presence of TraM in the inner membrane *in vivo* has been demonstrated using multicopy clones of *traM* (DiLaurenzio *et al.*, 1992). Therefore, the level of TraM in the membrane fraction of wild-type pOX38-Km, and in mutations in *traD* (F*lac traD8*) and *traG* (F*lac traG101*) was determined. We also demonstrated the effect of the *traM* mutation in pOX38*traM*K3 on the location of TraI. These results suggest that TraM is at best marginally associated with the membrane and is not affected by the absence of TraD or TraG while TraI is purely a cytoplasmic protein. TraD is an inner membrane protein (Panicker and Minkley, 1992) which was thought to target TraI to the membrane when cell fractionation was carried out using a multicopy plasmid overexpressing both proteins (Dash *et al.*, 1992). Certainly, the presence of a few copies of TraM and TraI bound to *oriT* of the single copy F plasmid might not be detectable by immunoblot analysis of membrane preparations. Another possibility is suggested by the results of Gordon *et al.* (1997) who reported that

F is found at precise locations within the cell. Using a LacI-GFP fusion that bound to a *lacO* repeat on F, they demonstrated that F is replicated at mid-cell and is abruptly transported to the ¼ and ¾ positions immediately prior to cell division. If conjugative replication is coordinated with vegetative replication, conjugation might be permissible only at certain times in the cell cycle and only at certain positions within the cell. The presence of recipient cells might trigger positioning of the relaxosome at the base of the transfer apparatus at specific times during this cycle. This step would constitute the "signal" that transfer should occur, a process long thought to involve TraM (Kingsman and Willetts, 1978). For instance, contact with the recipient cell might alter the conformation of a key protein in the transferosome such as TraD, allowing TraM bound to DNA within the relaxosome to interact with it. In rapidly growing cultures, mating efficiency decreases dramatically as cells leave early exponential phase suggesting that conjugation potential and cell growth are co-regulated (Frost and Manchak, 1998). The loss of transfer ability occurs at the same time as the cessation of *traM* transcription, suggesting once again that TraM has a role in this process that is not completely understood. These results suggest a model whereby the relaxosome moves to the membrane in response to mating pair stabilization thereby generating the elusive signal that triggers DNA synthesis and transfer. Hopefully, modern approaches to detecting proteins *in situ* will allow us to visualize where the transferosome/relaxosome complexes are located within the cell.

Chapter 4

Biochemical characterization of TraM in solution

Introduction

Previous characterizations of TraM using analytical ultracentrifugation (Di Laurenzio, 1992) did not give reliable results and suggested that TraM aggregated in solution. Even after denaturation in 4 M guanadine hydrochloride non-specific aggregates with molecular weights ranging from 30 000 to 120 000 Da were found. Sucrose density centrifugation (Di Laurenzio, 1992) and size exclusion chromatography (Penfold, 1995) however, suggested that TraM was a tetramer. Other problems such as non-reproducible association constants for various TraM binding sites warranted the repurification and reanalysis of TraM in order to determine its quaternary structure. The analysis of TraM from the F-like plasmid R1 (Kupelwieser *et al.*, 1998; Verdino *et al.*, 1999) also prompted similar analysis for F TraM since similarities and differences between these proteins have now been shown (Fekete and Frost, 2000).

Biochemical and biophysical characterization of proteins requires purified protein sample for optimal results. Contaminants can have confounding effects on results leading to incorrect conclusions. There are many purification procedures (Ausubel *et al.*, 1987), all of which must be tailored to obtain the best results for a given protein. These procedures take advantage of native characteristics of the desired protein to separate them from contaminants. For example, proteins can be separated based on charge, hydrophobicity, size, and specific binding characteristics (Protein Purification Handbook, Amersham Pharmacia biotech, Code No. 18-1132-29). To evaluate the physical and biochemical characteristics of TraM, purification to a high degree was required. Based on its characteristics, ion-exchange and DNA cellulose chromatography were chosen as the major steps in its purification.

Separation based on net charge is a common protein purification method. Mono Q® columns can be used to separate negatively charged proteins, and depending on the elution conditions can separate proteins based on degree of charge. Mono Q® columns use a matrix of MonoBeads®, a monodisperse hydrophobic resin with a small particle size (Amersham Pharmacia biotech, product notes, code no. 17-0546-01). The charged group on the matrix is $-CH_2-N^+(CH_3)_3$, and a variety of cationic buffers (or sodium chloride) can be used to elute bound proteins. The elution profile of the bound proteins depends on elution buffer, and protein charge, with more negatively charged proteins being retained on the column longer than less charged proteins.

Chromatography using DNA cellulose can be used to selectively bind proteins which bind DNA (Amersham Pharmacia biotech, product notes). The matrix is a cellulose fiber with native (double stranded) calf thymus DNA adsorbed to it. The exact forces which bind the DNA to the column are not known. The DNA is not covalently linked to the matrix, but is absorbed to the cellulose by drying. Most DNA binding proteins have an inherent non-specific DNA binding property, while non-DNA binding proteins do not. Using this non-specific binding property, unbound proteins can be washed from the column and bound proteins eluted by increasing sodium chloride concentrations, or by the addition of competing nucleic acids, biological inducers, or specific cofactors.

The separation of proteins by size is also a useful tool in purification. Gel filtration (or size-exclusion) chromatography requires concentrated samples and needs only a single buffer for elution so is often used as the final step in a purification procedure (Amersham Pharmacia biotech, Protein Purification Handbook). The porous nature of the gel filtration matrices allow small molecules to enter the beads which retards their elution, while larger molecules pass through the column and are eluted first. Superose 12®, commonly used in gel filtration experiments, is made of a cross-linked, agarose based medium with a very uniform bead size (Amersham Pharmacia biotech, product notes). These features allow the use of many types of solvents and ensure that flow rate is maintained. This procedure can also be used to quantitate the size of unknown multiprotein complexes (Amersham Pharmacia biotech, Gel Filtration Principles and Methods, cat #18-1022-18; Phillips *et al.*, 1996). Standard proteins of known sizes are run under the same conditions as the unknown protein and their molecular weights are plotted versus elution (taken as column volumes or fraction number) to construct a standard curve. Approximate molecular masses can then be read from the curve for unknown proteins or complexes. However, it should be noted that this technique is based on the cross sectional size of the molecule so can give misleading results if molecules are more or less dense than the standards used.

Mass spectrometry uses the differences in a molecules mass-to-charge ratio to identify its molecular weight (Howe *et al.*, 1981, pp. 1-26). Electrospray ionization mass spectrometry uses a needle to atomize samples which are accelerated using electrodes and then forced through a capillary (JEOL application note, MS71A). Molecules are then fed into a mass analyzer which separates molecules based on mass-to-charge ratios, and collected using a detector.

Amino acid analysis is often performed on purified proteins in order to determine amino acid content. This procedure usually involves hydrolysis of the protein in 6 M hydrochloric acid and heating to 110°C for 24 hours (Mathews and van Holde, 1990). This can be accelerated by heating the sample to a higher temperature. The amino acids are then separated by an ion exchange column and the eluent mixed with ninhydrin. The amino acids then fluoresce when excited at 570 nm and the identity of each amino acid is determined by the retention time on the column. Problems with this procedure include degradation of some amino acids, such as tryptophan, serine, threonine, and tyrosine, during the hydrolysis. Another problem which occurs is the hydrolysis of asparagine and glutamine to aspartic and glutamic acid, respectively. However, awareness of these complications can reduce the possibility of confusion upon analysis of the data.

Amino acid sequencing has made the identification of proteins a much faster process. Amino-terminal sequencing most often is done using the Edman degradation method discovered by Pehr Edman (Mathews and van Holde, 1990). This process uses the compound phenylisothiocyanate, which reacts with the terminal amino group of the peptide. The amino-terminal amino acid derivative is then cleaved from the peptide with an anhydrous acid, and isomerizes into a phenylthiohydantion (PTH) derivative of the amino acid. This derivative is then separated from the rest of the protein in ethyl acetate (in which the protein is insoluble), and run on a reversed phase HPLC column. The amino acid identity can then be determined by comparing the elution time of the derivative from the column to a set of PTH derivative amino acid standards run on the same column. The reaction can be repeated to determine the identity of the remaining amino acids. Sequencing can reliably process 40 to 60 residues using this method. Carboxyl-terminal sequencing can also be performed using carboxypeptidase Y, an enzyme which cleaves the last amino acid from a peptide (Mathews and van Holde, 1990). However, this procedure requires a specialized knowledge in order to determine in which order the amino acids are coming off of the peptide.

Analytical ultracentrifugation has become another popular method for the determination of molecular mass of a molecule or complex. This technique uses the principle of sedimentation equilibrium, which is when sedimentation and diffusion come to a steady state (Analytical Ultracentrifugation v.II, Beckman, 1993). This equilibrium is dependent on the buoyant molecular weight, angular velocity of the rotor, and temperature. The buoyant molecular weight is determined from the molecular weight of the molecule, its partial specific volume, and the density of the solvent. The value for partial specific volume is estimated based on amino acid sequence, but both of the latter two factors can have a profound effect on the determination of molecular weight if errors are made in their calculations. Most often samples are run at a variety of concentrations and at at least 2 speeds. Data can then be compared from a variety of conditions and conclusions made with greater surety.

Chemical crosslinking of proteins is another method of determining protein tertiary and quaternary structure. Many of the chemical compounds used can specifically react with different reactive groups such as amino, sulfhydryl, guanidino, indole, and carboxyl groups (Ji, 1983). This allows the spacial definition of these groups in a complex folded structure. Varying the linker arms of these chemical compounds can also allow for the approximation of distance between these reactive groups. Cleavable crosslinkers adds another dimension to the crosslinking methodology, and differentiates chemical crosslinking from chemical modification (Peters and Richards, 1977; Ji, 1983). A large selection of crosslinking reagents are now available (Pierce) which combine many of these features and a variety of others: chemical specificity, cross-bridge length, heterobifunctional (both ends react with different groups) or homobifunctional linkers, photochemically inducible linkers, cleavable linkers, and the ability of linkers to be radio-labeled. Crosslinked complexes can then be analyzed by SDS PAGE or by one of the other previously mentioned methods for determining size.

Results

The *traM* gene was cloned into the pT7-5 overexpression vector (Tabor and Richardson, 1985) and transformed into *E. coli* DH5α already containing pGP1-2, which encodes the gene for T7 RNA polymerase. Cells were induced by heat shock at 42^{O}C for 30 minutes (Figure 4.1, lanes 2 and 3) followed by lysis in a French Pressure Cell (Figure 4.1, lane 4). Unlysed cells were removed by centrifugation, and then membranes were pelleted in an ultracentrifuge (Figure 4.1, lane 5 and 6). Ammonium sulfate was added to 25% and precipitated proteins were collected in a centrifuge (Figure 4.1, lanes 7 and 8). The supernatant was desalted and passed over a MonoQ® HR 5/5 column. TraM eluted at a sodium chloride concentration of approximately 250 mM (Figure 4.1, lane 9). Fractions containing TraM were pooled and desalted and then passed over a 1 mL Native DNA Cellulose column. TraM eluted at a sodium chloride concentration of 300 mM (Figure 4.1, lane 10). Fractions containing TraM were again pooled and desalted and eluted from the desalting column in either 20 mM potassium phosphate or TED. These samples were then concentrated using a Amicon Stir Cell with a YM3 membrane (Figure 4.1, lane 11).

TraM suspended in 20 mM potassium phosphate was used for all protein analysis procedures. Mass Spectrometry was performed (Alberta Peptide Institute) and showed that TraM has a molecular weight of 14 376 ± 3 Da (Figure 4.2). This was surprising since the predicted molecular weight of TraM is 14 507 Da. Subtraction of the mass of a methionine amino acid (131.21 Da) gives 14 376 Da, suggesting that TraM does not contain its amino terminal methionine. Amino-terminal sequencing confirmed that TraM

Figure 4.1. Summary of steps involved in the purification of TraM. 0.1 O.D.$_{600}$ of cells or 20 µg of protein from each sample was separated on an 18% SDS polyacrylamide gel. Molecular weights of the markers are given on the sides in kDa. Gels were stained with Coomassie Brilliant Blue R. TraM is indicated with an arrow.

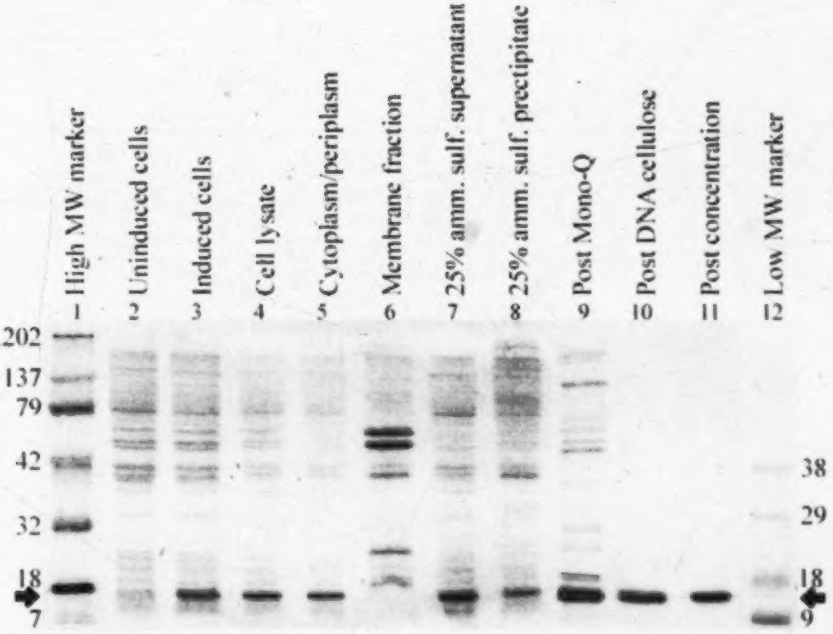
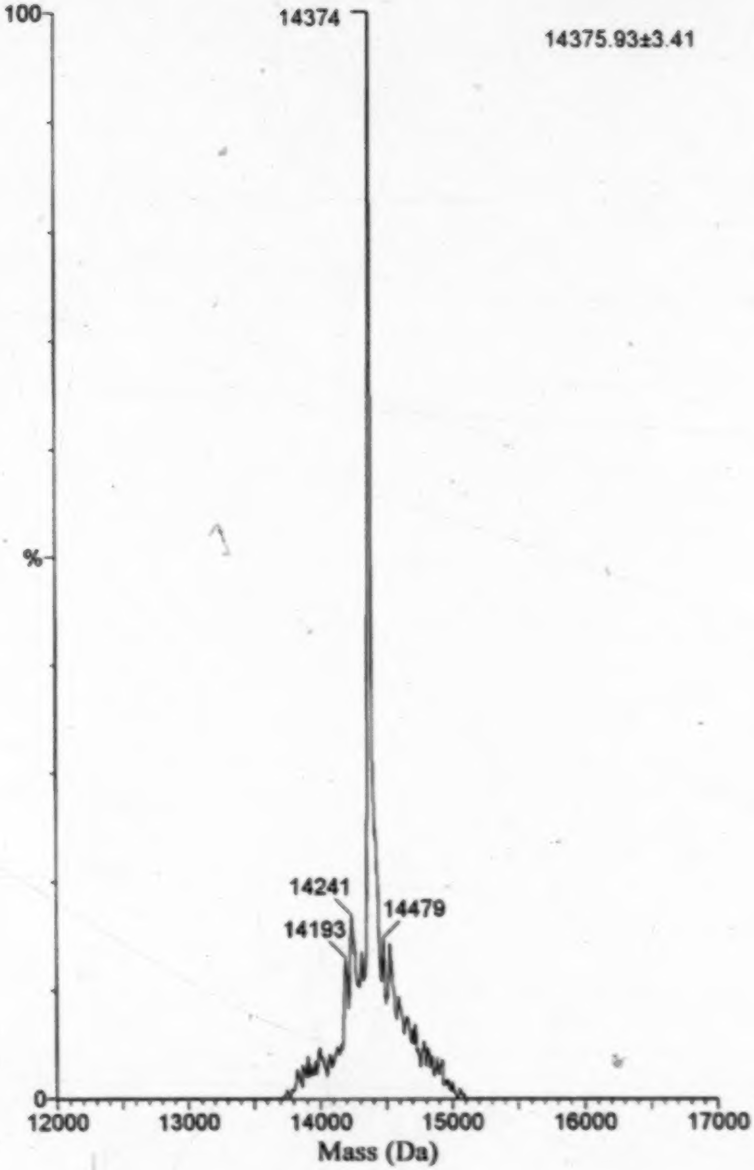

1 High MW marker
2 Uninduced cells
3 Induced cells
4 Cell lysate
5 Cytoplasm/periplasm
6 Membrane fraction
7 25% amm. sulf. supernatant
8 25% amm. sulf. prectipitate
9 Post Mono-Q
10 Post DNA cellulose
11 Post concentration
12 Low MW marker
202
137
79
42
32
18
7
38
29
18
9

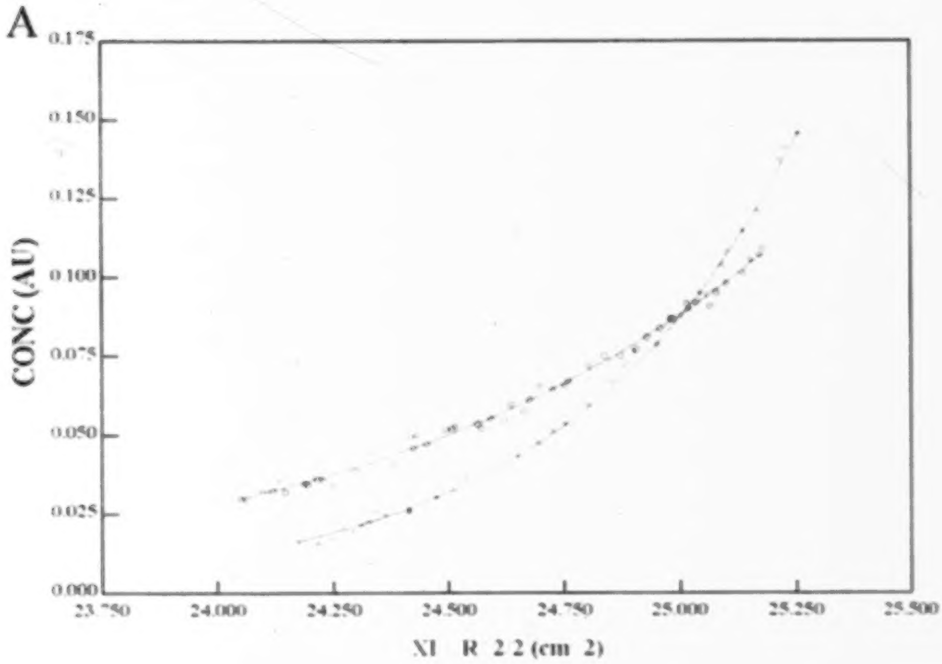
Figure 4.2. Mass spectrum of F TraM dissolved in TED. Masses of the major peaks are shown above in Da and the calculated molecular weight of TraM is shown at the top right.

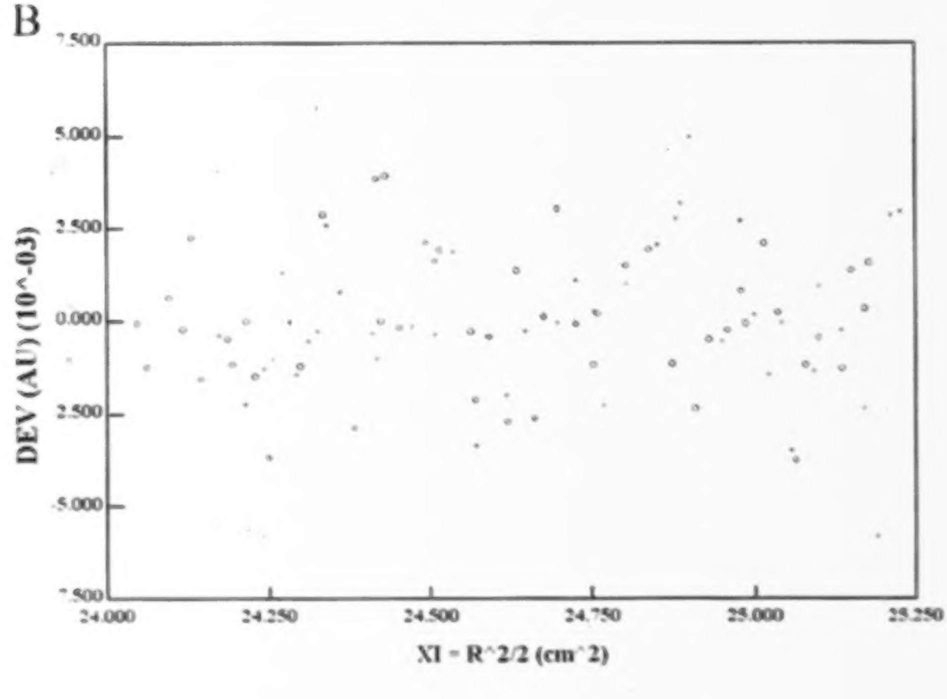

100
%
0
14374
14375.93±3.41
14241
14193
14479
12000
13000
14000
15000
16000
17000
Mass (Da)

did not contain its amino-terminal methionine, and showed that the following 10 amino acids were as predicted from the DNA sequence. TraM from the R1 plasmid (Verdino *et al*., 1999) has been shown to have its amino-terminal methionine cleaved, as is the case for F TraM. Amino acid analysis showed that the purified TraM was at a concentration of 3.7 mg/mL.

Analytical ultracentrifugation was performed by Les Hicks in the Dept. of Biochemistry, University of Alberta, on TraM in potassium phosphate buffer to determine the number and size of the species in solution. Concentrated TraM was diluted in 100 mM and 300 mM potassium chloride for analysis. Both samples were tested using two different rotor speeds (12 000 and 16 000 rpm) and three different protein concentrations (3, 7, and 13 μM). Figure 4.3a shows an example of the data plotted as concentration distributions at equilibrium (protein concentration versus radius of cell) at 2 speeds using 13 μM TraM in 100 mM potassium chloride. The upper line (circles) represents 12 000 rpm and the lower line (diamonds) represents 16 000 rpm. Data was then fitted to an equation for a single species model (lines in the graph). Non-ideality (from aggregation for example) would cause the graph to be more linear. If two species were present then the data would have more of a curve. Residual values (variances from the model) were plotted on a Residuals plot (Figure 4.3b, deviation or variance versus radius of cell). If the data fits well to the model then residual values should be found evenly distributed around the zero value. Non-ideality would cause the variance values to appear in a frown pattern. If two species were present (when using a single species model) the variance values would appear in a smile pattern along the zero value. The square root of the variance is another

Figure 4.3. A. Graph showing concentration distribution at equilibrium plotted as protein concentration versus the radius of the sample cell. TraM at 13 μM in 100 mM potassium chloride is shown at two speeds, 12 000 (circles) and 16 000 rpm (diamonds). B. Residuals plot for the data from A fit to a single species model. The zero value on the Y-axis represents no deviation from the model. Variations from the single species model for each value in A is shown as circles (for the 12 000 rpm run) and diamonds (for the 16 000 rpm run).

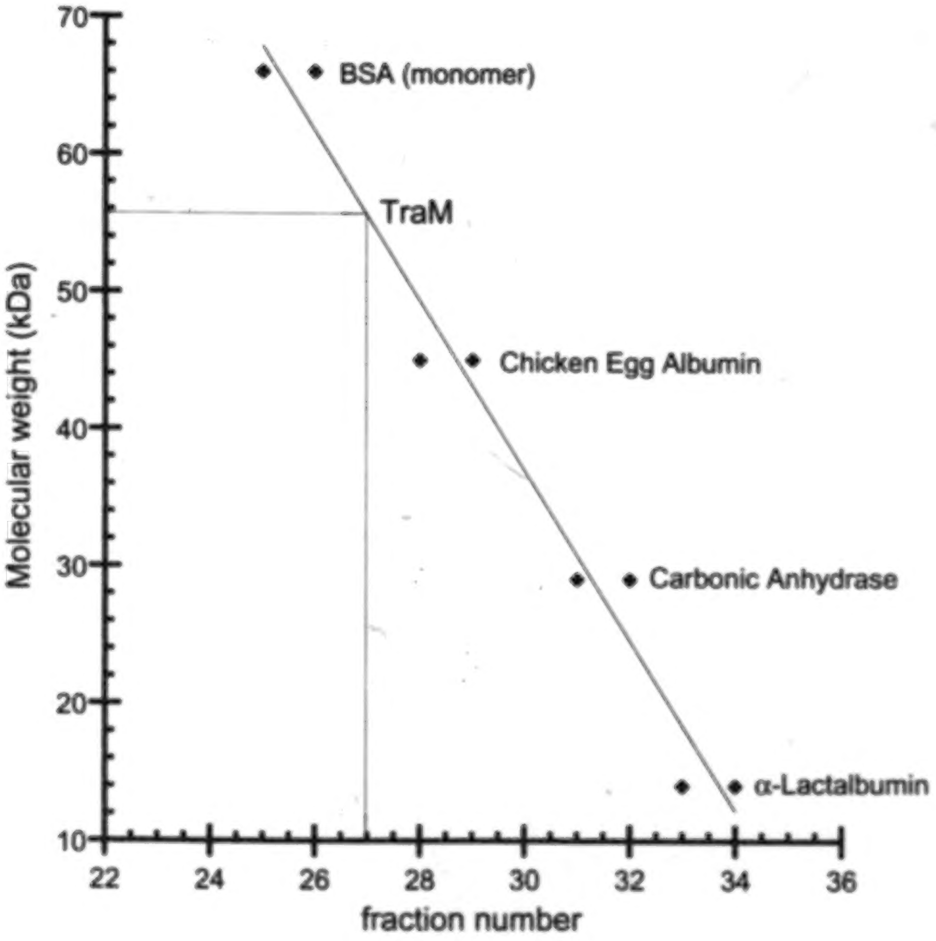

A
CONC (AU)
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
23.750
24.000
24.250
24.500
24.750
25.000
25.250
25.500
XI = R^2/2 (cm^2)


B
DEV (AU) (10^-03)
7.500
5.000
2.500
0.000
-2.500
-5.000
-7.500
24.000
24.250
24.500
24.750
25.000
25.250
XI = R^2/2 (cm^2)

way to test the model's fit. The square root of the variance should be less than $2X10^{-2}$ for a good fit. The square root of the variance for TraM in this single species model was calculated to be $2.3X10^{-3}$. Since the data fits well to a single species model it suggests that TraM is a single species at these concentrations. Using the equation from the single species model to calculate the molecular weight gave 66 014 Da. This suggests that there are 4.6 molecules of TraM per multimer, which suggests that TraM most likely exists as a tetramer. However, it should be noted that the partial specific volume of the molecule is used in the equation to calculate molecular weight. This value is usually estimated based on a prediction from the protein composition. However, errors in this value are magnified three to five fold (Self-Associating Systems in the Analytical Ultracentrifuge, Beckman, 1993). If the partial specific volume is overestimated by 1% (molecule is denser than estimate) the molecular weight will appear to be 3-5% higher than it really is. This value can be calculated by doing densitometry or analytical ultracentrifugation using different solvents. However, densitometry requires specialized instruments, and ultracentrifugation with different solvents requires parallel runs using H_2O and D_2O. Therefore, these experiments were not pursued.

Size exclusion chromatography was performed on 400 μg of TraM in 20 mM potassium phosphate using a Superose 12® column. Standards were also run on the same column using the same conditions and a standard curve was made for the elution of the standards (Figure 4.4). TraM eluted in fractions 26 to 28 and a line representing fraction 27 is drawn on the curve. These fractions correspond to molecules having a molecular weight between 50 000 and 62 000 Da. Taking the center of these values gives a molecular

Figure 4.4. Standard curve constructed using a set of protein standards in size exclusion chromatography experiments. Molecular weight of protein standards is graphed versus fraction number from chromatographic runs using a Superose 12® column. Two points are given for each protein standard, representing the first and last fraction in which that protein was found. A vertical line designates the middle fraction in which TraM was found, with a horizontal line showing the molecular weight that fraction represents.

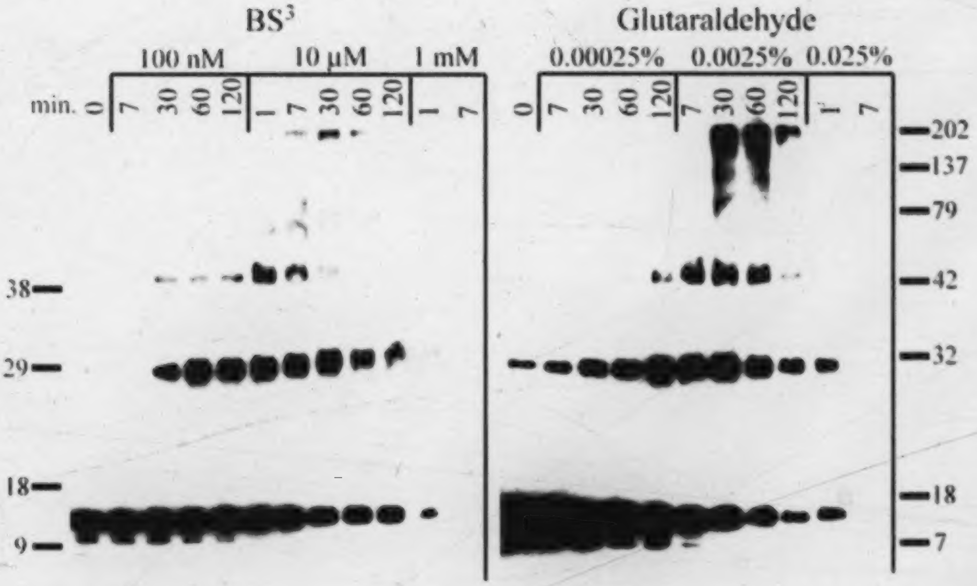

70
60
50
40
30
20
10
Molecular weight (kDa)
22
24
26
28
30
32
34
36
fraction number
BSA (monomer)
TraM
Chicken Egg Albumin
Carbonic Anhydrase
α-Lactalbumin

weight of 56 000 Da which is very close to the predicted molecular weight of 57 504 for a TraM tetramer. This suggests that under these conditions TraM is found as a tetramer.

Chemical crosslinking using glutaraldyhyde, BS^3, and DSP (Pierce) were performed on TraM in solution and bound to DNA. All three reagents crosslink primary amines located on the amino terminus of the protein and ε-amines of surface lysines (Peters and Richards, 1977; Pierce). These reagents were used since TraM contains 11 lysines. Glutaraldehyde also reacts to a lesser degree with the side chains from cysteine, histidine and tyrosine (Ji, 1983). The crosslinkers differ in their linking arm size and this is used as a common variable in crosslinking experiments (Rousseau *et al.*, 1996). BS^3 has a crosslinking arm of 11.4 angstroms and DSP an arm of 12 angstroms. Polymerization of glutaraldehyde makes estimating the crosslinking arm size of glutaraldehyde impossible, however, this arm can be as few as 4 carbon atoms long (Peters and Richards, 1977). TraM was used at a concentration of 0.5 μM and, after the reactions were quenched, was run on 15% SDS denaturing polyacrylamide gels. After transfer immunoblots were performed on the membranes using anti-TraM antibodies. Examples of these experiments are shown in Figure 4.5. DSP gave the same pattern as BS^3 so it was not used further. BS^3 was used at concentrations of 100 nM to 1 mM and glutaraldehyde was used at concentrations of 0.00025% to 25% (only 0.00025% to 0.025% are shown in Figure 4.5). Crosslinking data suggests that both BS^3 and glutaraldehyde cause the formation of dimers (approximately 29 000 Da) and tetramers or trimers (approximately 50 000 Da). Dimers of TraM were occasionally seen in the absence of crosslinking reagent (0 min. in glutaraldehyde), and seemed to depend on the amount of denaturation.

Figure 4.5. Immunoblots of TraM run in 15% SDS PAGE after crosslinking with BS^3 and glutaraldehyde. Final concentrations of each of the crosslinkers is shown at the top, as is the time, in minutes, for each crosslinking reaction. The 0 values represent no crosslinking reagent added. The position of the Low (left side) and High (right side) molecular weight markers is also shown in kDa.

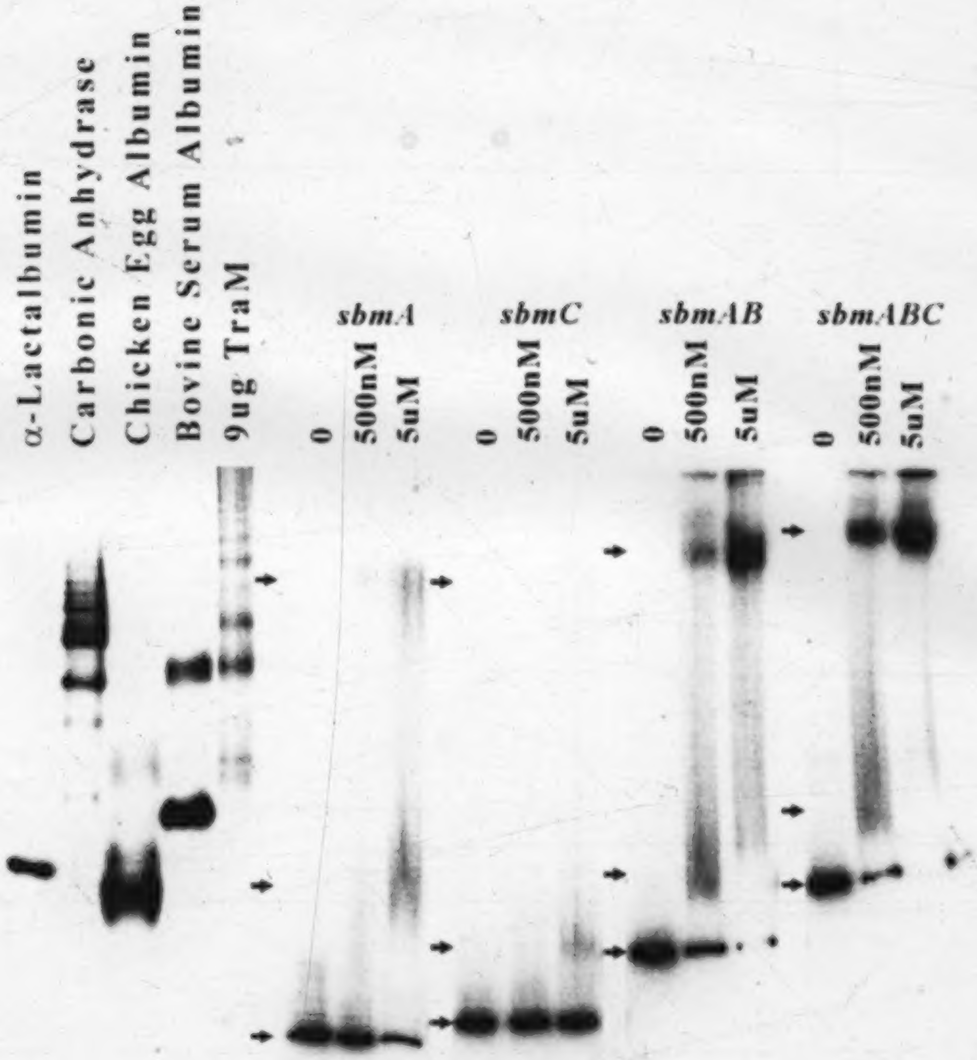
6


BS3
100 nM
10 μM
1 mM
Glutaraldehyde
0.00025%
0.0025%
0.025%
min.
0
7
30
60
120
1
7
30
60
120
1
7
0
7
30
60
120
7
30
60
120
1
7
38
29
18
9
202
137
79
42
32
18
7

The increase in dimers upon the addition of crosslinking reagent confirms that dimers were being chemically linked together. Larger species were seen at higher crosslinker concentrations and this is likely to be non-specific crosslinking. At very high crosslinking concentrations no protein was seen and may not have entered the gel due to the large size of the crosslinked aggregate. Binding of TraM to any of its binding sites (*sbmA*, *sbmC*, *sbmAB*, or *sbmABC*) followed by crosslinking did not have any effect on the pattern seen, suggesting that DNA binding does not have an effect on the conformation of the crosslinking sites. This is different from what has been seen in other studies where DNA binding can alter crosslinking efficiency (Kersten *et al.*, 1995; Phillips *et al.*, 1996). To determine if glutaraldehyde and BS^3 crosslink the same regions of TraM, both reagents were used in the same reaction. However, only large aggregates of TraM were seen instead of the predicted tetramers.

To determine the size of the multimers bound to the various TraM binding sites non-denaturing gel electrophoresis was performed. These experiments work on the principle that larger molecules will have different mobility patterns compared to small molecules when run on a variety of gel concentrations (Orchard and May, 1993). This means that in respect to their migrating speed, larger molecules are more affected by increases in gel concentration than are small molecules. Radioactively labeled DNA (as used in retardations) was used in these experiments in order to be able to identify the retarded multimers. TraM bound to *sbmA*, *sbmC*, *sbmAB*, and *sbmABC* were run with standards of Bovine Serum Albumin, Chicken Egg Albumin, Carbonic Anhydrase and α-Lactalbumin and is shown in Figure 4.6. TraM at 500 nM and 5 μM was used for

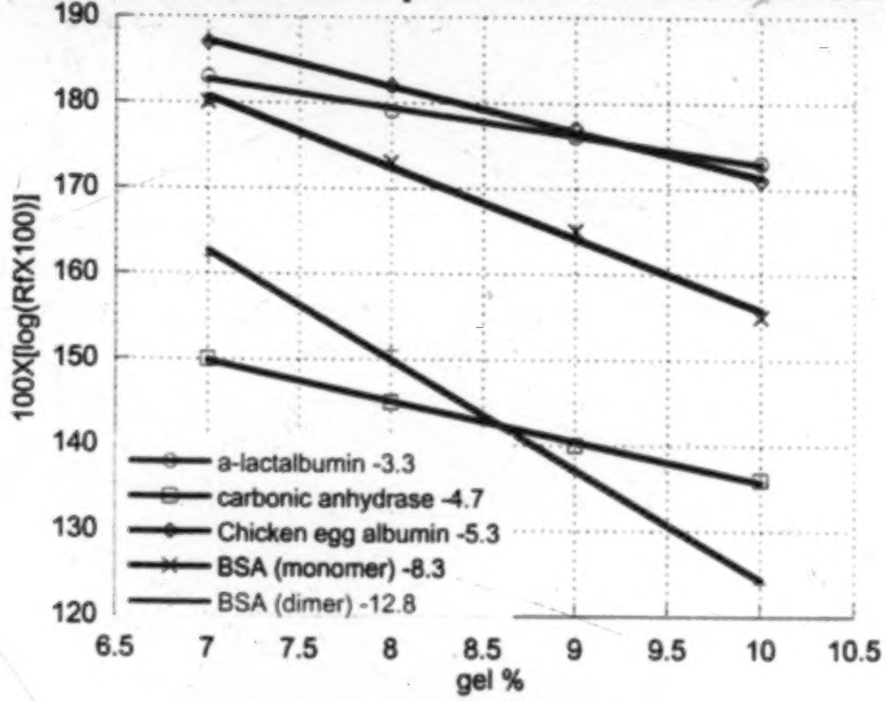
Figure 4.6. Example of the series of non-denaturing polyacrylamide gels used to determine the size of protein complexes bound to TraM binding sites. Protein standards and retarded TraM binding sites are shown on a 9% gel. Concentrations of TraM used in the retardation reactions are shown above each lane. Position of the unbound, lower, and upper complexes are shown with arrows.


α-Lactalbumin
Carbonic Anhydrase
Chicken Egg Albumin
Bovine Serum Albumin
9ug TraM
sbmA
0
500nM
5uM
sbmC
0
500nM
5uM
sbmAB
0
500nM
5uM
sbmABC
0
500nM
5uM

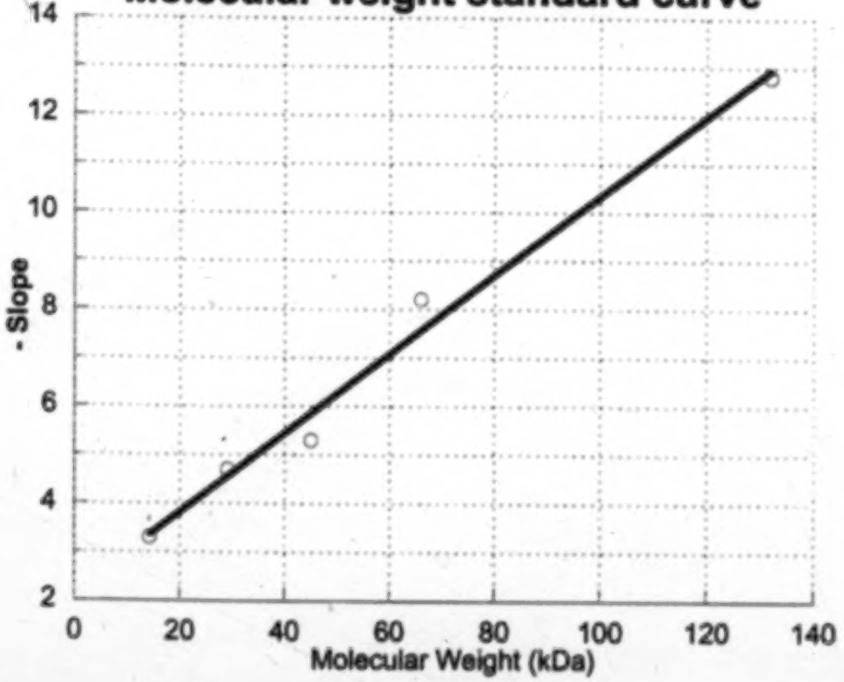
retardations of all of the binding sites. A 9% polyacrylamide gel is shown, however, all samples were run on 7, 8, 9, and 10% polyacrylamide gels. The mobility of all bands (protein standards, unbound and bound DNA) were measured from the top of the separating gel and compared to the mobility of the Bromophenol Blue dye to give their relative mobilities (R_f = distance of protein migration/distance of BPB dye migration). These were then entered into the equation $100[\log(R_f \times 100)]$ and the results plotted versus gel percentage (Figure 4.7). The negative slopes of the protein standards were then plotted versus molecular weight to make a standard curve (Figure 4.7, below). Relative mobilities of all other bands (free DNA and DNA:protein complexes) were placed into the above equation and plotted versus gel percentage (Figure 4.8). Negative slopes from these plots were taken and their molecular weight calculated using the standard curve. The size of the protein multimer bound to the DNA was then calculated by subtracting the molecular weight of the unbound DNA from the bound DNA. The DNA fragments used in these experiments had to be short since long pieces of DNA were found to mask the effect of differences seen when altering gel concentrations. For this reason *sbmA* was 100 bp, *sbmC* was 100 bp, *sbmAB* was 160 bp, and *sbmABC* was 210 bp.

Data from these experiments is summarized in Table 4.1. For *sbmA* TraM bound to the lower band corresponds to 36 000 Da and the upper retarded band to 56 000 Da. These sizes may correspond to a TraM dimer binding initially followed by tetramerization on the DNA. For *sbmC*, TraM bound to the lower band corresponds to 12 000 Da and the upper to 48 000 Da. TraM bound to the lower band of *sbmAB* corresponds to 10 000 Da

Figure 4.7. Graph showing the effects of gel concentration on the relative mobility of protein standards (upper panel). The slopes of these curves (see figure) were then plotted versus molecular weight to generate a standard curve (lower panel) which was used to determine the molecular weight of unknown molecules.



Relative mobilities of proteins used as standards



Molecular weight standard curve

Figure 4.8. Graphs showing the effect of gel concentration on the relative mobility of TraM binding sites (bound and unbound). The slope of these graphs (see figure) was used to determine the molecular weight of each species using the standard curve (Figure 4.7).

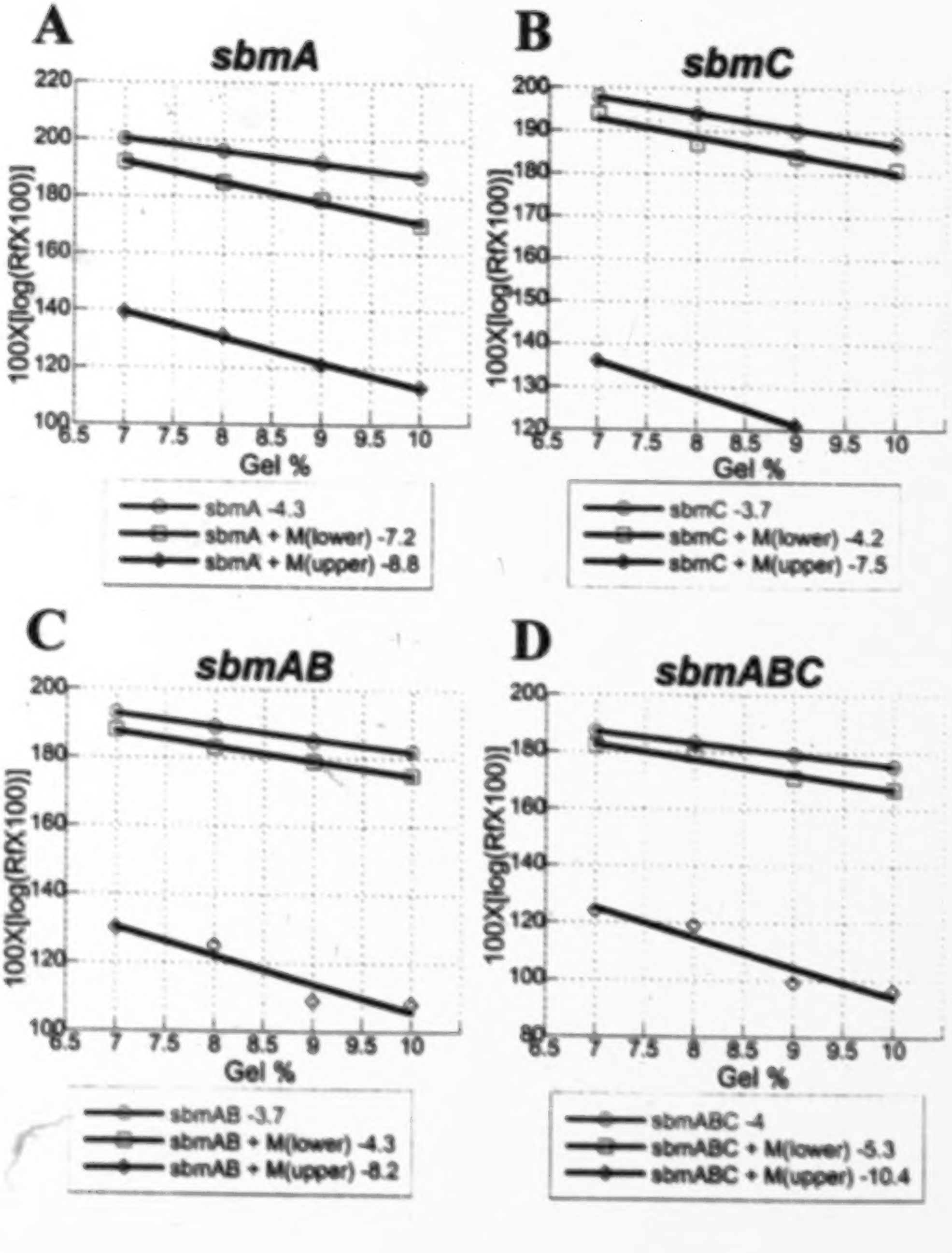

A
sbmA
100X[log(RfX100)]
220
200
180
160
140
120
100
6.5
7
7.5
8
8.5
9
9.5
10
Gel %
sbmA -4.3
sbmA + M(lower) -7.2
sbmA + M(upper) -8.8
B
sbmC
200
190
180
170
160
150
140
130
120
sbmC -3.7
sbmC + M(lower) -4.2
sbmC + M(upper) -7.5
C
sbmAB
sbmAB -3.7
sbmAB + M(lower) -4.3
sbmAB + M(upper) -8.2
D
sbmABC
80
sbmABC -4
sbmABC + M(lower) -5.3
sbmABC + M(upper) -10.4

Table 4.1

Size of protein complex bound to TraM binding sites

Fragment	Calculated Size of lower band	Proposed Multimer	Calculated Size of upper band	Proposed Multimer
sbmA	36 kDa	Dimer	56 kDa	Tetramer
sbmC	12 kDa	monomer	48 kDa	Tetramer
sbmAB	10 kDa	monomer	50 kDa	Tetramer
sbmABC	15 kDa	monomer	73 kDa	Tetramer +Dimer

and the upper to 50 000 Da. Whether the lower bands represents a monomer is not certain, however, the size of TraM bound to the upper band suggests that this may be a tetramer. Calculated values for TraM bound to the lower and upper bands of *sbmABC* are 15 000 Da and 73 000 Da respectively. Whether the lower band corresponds to a monomer is again not certain, however, the upper band may correspond to a tetramer and a dimer of TraM bound to the DNA. Since large fragments of DNA can cause an underestimation of the size of the protein bound to the DNA, the lower bands of *sbmAB* and *sbmABC* may actually represent dimers, not monomers. In these cases the retardation effect of the dimer on the larger DNA molecule may be quenched by the lack of a substantial increase in the molecular weight of the DNA complex.

Discussion

Overexpression of F TraM was performed using the T7 RNA polymerase and pT7 plasmid system. TraM was purified to a high degree using anion ion exchange and native DNA-cellulose chromatography columns. This protocol allowed mass spectroscopy, amino-terminal sequencing, amino acid analysis, and analytical ultracentrifugation to be performed on the protein.

Amino-terminal sequencing and mass spectrometry demonstrated that TraM has a molecular weight of 14 376 ±3 Da and that the amino-terminal methionine was not present. A similar finding was found with TraM from the plasmid R1 (Verdino *et al.*, 1999). This phenomenon is common in *E. coli*, where approximately half of proteins do not contain their amino-terminal methionine. This occurs when the amino-terminal N-formyl-methionine is deformylated by a deformylase and the amino-terminal methionine is removed by an aminopeptidase (Lewin, 1994, pp. 179-182). Amino acid analysis showed that the protein concentration was higher than what was found using Lowry assays (Ausubel *et al.*, 1987), however, this may be due to the fact that the Lowry assay relies on the radical groups of tyrosine, tryptophan, and cysteine to react with Folin reagent. Since these amino acids only compose 3.2% of TraM as opposed to the 7.4% found in nature (Mathews and Van Holde, 1990, pp. 137) this is not unexpected.

Data from analytical ultracentrifugation of the purified protein preparation suggested that TraM exists as a macromolecule of approximately 66 000 Da in solution. This most likely correlates with a tetramer whose predicted molecular weight would be approximately 57 000 Da. These differences could be due to the fact that the tetramer has a smaller partial specific volume than predicted. The close adherence to the single species model suggests that TraM is present only as a tetramer in solution at the concentrations tested. This is in contrast to what was found before by Laura Di Laurenzio (1992) for F TraM when analytical ultracentrifugation was performed on TraM in TED buffer. In those studies TraM was found to aggregate and displayed molecular weight ranging from 40 000 to 120 000 Da. After denaturation in 4 M guanidine hydrochloride and its removal by dialysis aggregates were still found with calculated weights ranging from 30 000 to 90 000 Da, suggested to represent TraM as a mixture of multimers ranging from dimers to hexamers. The improved performance of TraM during ultracentrifugation may be a result of the difference in quality of the protein which may have resulted from some minor differences in the two purification procedures. However, using sucrose gradient centrifugation, TraM previously was found to be a tetramer (Di Laurenzio, 1992). When compared to known size standards, TraM had a molecular weight of approximately 60 000 Dalton which was found to be reproducible over a pH range of 6.4 to 8.8.

Using size exclusion chromatography TraM is suggested to have a molecular weight of approximately 56 000 Da. This is in close agreement with the proposed molecular weight of a TraM tetramer of 57 504 Da. Sonya Penfold (1995) also performed size exclusion chromatography on TraM using bovine serum albumin and carbonic anhydrase as size standards. However, these experiments were done on a crude extract and not on a purified protein preparation. In those experiments TraM was found to have a molecular weight of approximately 50 000 Da and therefore, was also suggested to be mainly tetramers.

Chemical crosslinking of TraM was performed to determine the multimeric state of TraM in solution and to differentiate these interactions. Three different crosslinkers were used which allowed the crosslinking arms to vary from approximately 4 carbon atoms to 12 angstroms. It was hoped that this method would be able to separate the dimerization and tetramerization domains by purely physical distance with each reagent crosslinking separate domains. However, TraM treated with BS^3 and glutaraldehyde formed both dimers and tetramers (or trimers). This was disappointing but may be due to the fact that both react with the same chemical groups (ε-amines). Another complicating factor is the polymerization characteristics of glutaraldehyde which can make crosslinking arms of almost any size (Peters and Richards, 1977). The results did show that as crosslinking concentration or reaction time increased, more dimers, tetramers (or trimers), and finally larger aggregates were seen. The absence of smears, or non-specific aggregates, suggests that there are specific dimerization and tetramerization domains which are being linked. The presence of dimers without crosslinking was also seen and seemed to depend on the extent of denaturation in SDS before running on a gel. This suggests that the domain which gives these dimers is extremely stable.

Crosslinking carried out when protein is bound to DNA can produce a different crosslinking profiles. An example of this is the T4 RegA protein which gave dimers, trimers and higher oligomers in solution, but when bound to gene 44RE RNA gave mainly monomers (Phillips *et al.*, 1996). Another example is the mammalian retinoid X receptor (RXR) which gave monomers, dimers and tetramers during crosslinking reactions in solution (Kersten *et al.*, 1995). However, when bound to the RXR consensus response element DR-1, mainly dimers and monomers were found. In these cases the absence of the higher order multimers when bound to their target sites is thought to be due to the shielding of protein-protein interacting domains (RegA) or due to the inability of the protein to self aggregate (RXR). Another possibility is that the protein assumes a different conformation when bound to the DNA. Examples of this is R1 TraM which increases in α-helical content from 48% to 58% upon binding DNA (Verdino, P., *et al.*, 1999). This change in conformation could then alter the interacting faces of the protein. When crosslinking of TraM bound to DNA (*sbmA*, *sbmC*, *sbmAB*, *sbmABC*) was performed, no difference in the crosslinking profile was seen. This suggests that the dimerization and tetramerization domains are stable and are not dramatically altered upon binding to the TraM binding sites.

Non-denaturing gel electrophoresis was used to determine the number of TraM subunits bound to each of the cloned binding sites. TraM bound to radioactively labeled DNA was electrophoresed and the position of the bands was measured and compared to standards run in the same gel. From this data it appears that initially a dimer binds to *sbmA* followed by binding of another dimer. For *sbmC* and *sbmAB* it appears that a monomer is bound in the lower band and a tetramer bound in the final complex. For *sbmABC* it appears that a monomer binds followed by binding of more TraM resulting in a hexamer bound to the 3 binding sites. It seems unlikely that in the initial stages of binding a monomer would bind to *sbmAB* but a dimer bind to *sbmA*. Since *sbmAB* and *sbmABC* are larger DNA fragments (160 and 210 bp respectively), smaller protein complexes, such as a dimer, may be underestimated and may appear to be a monomer. Larger protein complexes such as tetramers and hexamers appear to be able to overcome this, but are still affected making the determination of the exact protein size impossible. For example, the final complex when bound to *sbmABC* is estimated to be 73 000 Da which would be approximately 5.1 TraM subunits. Since this is most likely an underestimation, a hexamer is proposed to be bound to this site.

Chapter 5

Identification and Analysis of the TraM multimerization domains using the yeast two-hybrid system

Introduction

The discovery that TraM exists as tetramers in solution, and as dimers and tetramers when bound to its DNA binding sites (Di Laurenzio, 1992; previous chapter) indicates that there are domains within TraM responsible for protein:protein interaction. Identifying these domains can facilitate the proposal of models for tetramerization and consequently, the mechanisms of DNA binding. For example, the proposal of a DNA looping model between LacR binding sites was elucidated by the identification of the protein multimerization domains (reviewed in Friedman *et al.*, 1995). Identifying the domains responsible for protein:protein interaction in TraM may lead to the proposal of a model for tetramerization in solution. This may lead to a model for how TraM binds single sites, such as *sbmA*, and multiple sites, such as *sbmABC*.

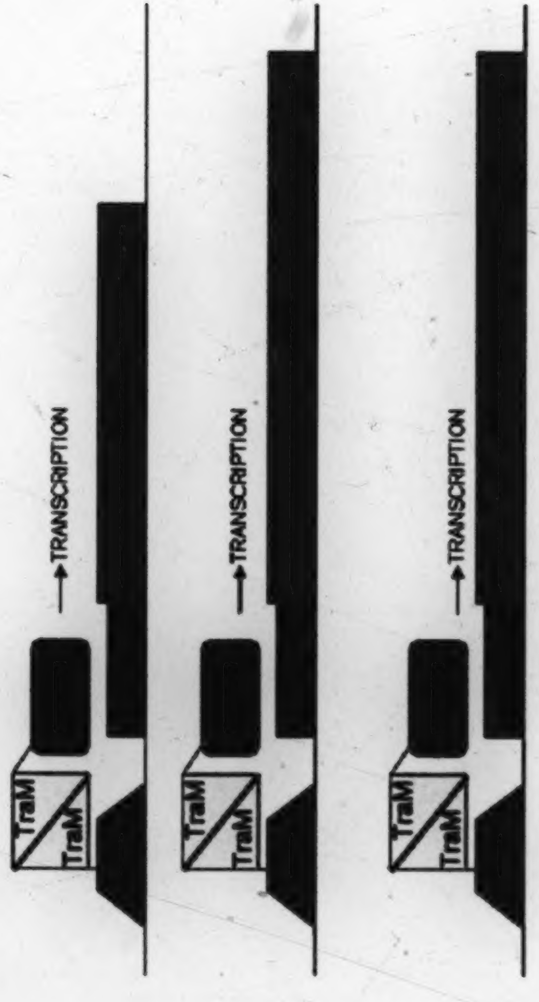
Examining how proteins interact in multiprotein complexes is important in order to understand the function and mechanism of action of each protein. Determining which regions of the proteins are required for this interaction can lead to the proposition of models for how each protein fits into these complexes. Many types of strategies exist which can define these interacting regions. However, most of these experiments are performed *in vitro* where purification or chemical treatments may affect protein structure. One example of this type of assay is the coprecipitation, or immunoprecipitation of protein complexes using antibodies (Ausubel *et al.*, 1987). However, this requires the purification of large amounts of protein and the availability of antibodies to these proteins. This technique is often plagued with the identification of proteins which do not interact specifically with the protein of choice and requires enough protein to be coprecipitated to be identified by some means such as amino acid sequencing. Another example is Far Western analysis or Affinity blotting (Phizicky and Fields, 1995) where proteins are transferred to a membrane, probed with a selected protein, and then probed using antibodies to the probing protein. However, this method requires the knowledge of which proteins have been transferred to the membrane, and as above the purification of large amounts of protein and the availability of their antibodies. The use of size exclusion or gel filtration chromatography (Ausubel *et al.*, 1987) to isolate large complexes has also been used. These techniques rely on strong interactions to hold the complexes together during cell lysis and multiple purification steps. Another downfall of these types of experiments is that they do not quantitate the strength of these interactions and a minimal decrease in the strength of interaction can yield negative results.

The yeast two-hybrid system (Y2HS; Fields and Song, 1989) allows determination of protein:protein interaction *in vivo* where proteins are more likely to be in their native conformation. This dispels many of the problems associated with *in vitro* methods. This system uses the *Saccharomyces cerevisiae* GAL4 protein, which is a DNA binding transcriptional activator. The DNA binding domain resides in the amino-terminal 147 amino acids of GAL4 and does not activate transcription (Keegan *et al.*, 1986). The transcriptional activating domain of GAL4 is located in the carboxyl-terminal 113 amino acids of GAL4 and does not activate transcription alone because it does not localize to the DNA. Genes of interest (such as *traM*) can be fused to the carboxyl-termini of the activating domain and the DNA binding domain and coexpressed in yeast. The strain

PJ69-4A (James *et al.*, 1996) used here is auxotrophic for leucine, and tryptophan. Genes complementing the auxotrophy are placed onto the plasmids containing the GAL4 fusions as a method of plasmid selection and stability. Each of the fusion proteins is expressed at high levels from the constitutive ADH1 promoter. Both the DNA binding domain and transcriptional activating domain fusions are then targeted to the nucleus. The DNA binding domain contains intrinsic nuclear localization signals (Silver *et al.*, 1984), and the transcriptional activating domain contains the SV40 T-antigen nuclear localization signal fused upstream of the activating domain (Chien *et al.*, 1991). When interactions occur between the fusions, the transcriptional activating domain is localized to the DNA binding domain, which binds sequences located upstream from the reporter genes. Therefore, interactions between the fusion proteins are detected by the activation of transcription from the reporter genes. PJ69-4A has three reporter genes, *HIS3*, *ADE2*, and the *E. coli lacZ* gene, cloned into the chromosome (Figure 5.1). This strain is auxotrophic for histidine and adenine synthesis, so will not grow if these amino acids are not provided in the media or if the reporter genes are not activated. The *lacZ* reporter gene is used to test the strength of protein interactions using β-galactosidase assays. Each reporter gene was placed downstream of a different promoter (each containing the GAL4 binding site) in order to reduce the chances of induction of all reporter genes by false positives (see below).

The yeast two-hybrid system can be used to screen large libraries of genes to find proteins interacting with a selected protein (James *et al.*, 1996). Since the Y2HS uses a positive selection approach, the identification of proteins which interact is simplified by

Figure 5.1. A diagrammatic representation of the reporter genes in PJ69-4A. The reporter genes, shown in green, are cloned into the chromosome (not shown to scale). The Gal promoters are shown in purple and their designations are inside each box. TraM is shown in yellow and the GAL4 activating and binding domains are designated as GAL4-AD and GAL4-BD respectively and are shown in red.


TRANSCRIPTION
TraM
TraM
TRANSCRIPTION
TraM
TraM
TRANSCRIPTION
TraM
TraM

growth on selective plates. Once interactions are identified, plasmid clones containing the putative positive genes can be transformed into *E. coli* and isolated. Candidates can then be identified by sequencing the DNA cloned into the plasmid. The advantage of this system is that positive clones are then available for sequencing, retransformation, and storage without worry of loss or degradation. The Y2HS has also been utilized to define domains of proteins which interact (Leanna and Hannink, 1996; Kalpana and Goff, 1993; Takacs *et al.*, 1993; Luban *et al.*, 1993). In these experiments segments of proteins which are known to interact are cloned and tested for their ability to interact with either full length proteins, or specific regions of other proteins.

Since its inception, modifications of the Y2HS have been developed. For example, a "reverse" yeast two-hybrid system has been used to look at the interactions of c-Rel and p40 proteins (Leanna and Hannink, 1996). In this system the yeast strain used is resistant to cyclohexamide due to a mutation in the *CYH2* gene. This mutation changes a glutamine at position 37 to a glutamic acid giving cyclohexamide resistance. However, sensitivity to the drug is still dominant (Sikorski and Boeke, 1991). A wild type copy of the *CYH2* gene is then placed under the control of the Gal1 promoter and interaction of fused activating domain and binding domain proteins makes the yeast non-viable on plates containing cyclohexamide. This system allows for the screening of mutations which inactivate the interaction between fused proteins. Another variation of the Y2HS is the yeast one-hybrid system (Y1HS). Bush *et al.* (1996) used this system to define the estradiol binding sites in the human estrogen receptor (HEGO). When the human estrogen receptor binds estradiol it possesses transcriptional activating activity, however, when estradiol is not bound, no transcriptional activation activity is present. HEGO was fused to the GAL4 DNA binding domain and induction of reporter genes was screened for in the presence of estradiol. Another variation is the yeast three-hybrid system (Y3HS). The basis of this system requires a three protein complex to be formed in order for reporter gene activation. For example one protein is fused to a DNA binding domain of GAL4, another is freely expressed which interacts with the DNA binding domain fusion, and a third protein is fused to the transcriptional activating domain of GAL4. Only when the two fusion proteins are linked via the third protein are reporter genes activated. This system has been used to examine the interactions of three proteins (Licitra and Liu, 1996) or with an RNA molecule linking the two fusion proteins (Wang *et al*., 1996). Many other modifications of the Y3HS are prevalent but are based on these same mechanisms. Another variation is the bacterial one- and two-hybrid systems (reviewed in Hu *et al*., 2000) which have studied peptides as small as 10 amino acids and proteins as large as 1179 amino acids. The bacterial one-hybrid system uses the DNA binding portions of the Lambda repressor fused to a protein of choice. If the selected fusion protein multimerizes then the repressor is able to dimerize and cooperatively bind to its binding sites. This can be detected by phage immunity or repression of a reporter gene depending on which test system is used. Without interaction, no binding occurs. Many of the bacterial two-hybrid systems use chimeric promoters consisting of binding sites for proteins like the Lambda repressor, 434 repressor, or LexA. Selected proteins can be fused to two of these proteins and interaction between the fusion proteins leads to cooperative binding or looping of the DNA at the chimeric promoter which results in repression of a reporter gene. Tetramerization of proteins can even be tested using two

Lambda operators upstream of a reporter gene. Dimerization of the fusion proteins causes poor repression, however, tetramerization fully represses the reporter gene.

The Y2HS has also been used to identify proteins from many different systems such as bacterial NifA and NifL (Lei *et al.*, 1999), yeast proteins such as CDK2/p21Cip1 (Cayrol *et al.*, 1997), human proteins such as p53 (Iwabuchi *et al.*, 1993), and viral proteins such as HIV Gag proteins (Franke *et al.*, 1994). The Y2HS has also been used to identify many types of proteins such as transcriptional regulators like the human autoimmune regulator (AIRE; Pitkanen *et al.*, 2000), DNA binding proteins like Jun and Fos (Chevray and Nathans, 1992) and c-Myc and Max (Kato *et al.*, 1992), cytoskeletal proteins like human 4.1R (Hou *et al.*, 2000), membrane proteins like the neuronal voltage-dependent sodium channel (VDSC; Mori *et al.*, 2000), and cytoplasmic proteins such as the mouse aromatic hydrocarbon receptor (AhR; Ma and Whitlock, 1997).

Although inherent problems do exist with the use of the Y2HS, proper controls can minimize them. Acidic sequences are known to activate transcription in yeast (Ma and Ptashne, 1987a,b). However, testing fusions by themselves and with control plasmids can eliminate these false positives. If this problem persists, these proteins can be fused to the transcriptional activating domain of GAL4 and the other protein or library fused to the DNA binding domain. Similarly, proteins possessing DNA binding activities can be fused to the DNA binding domain of GAL4 and screens can be performed using the transcriptional activating domain. Many of the false positives that are isolated usually affect only one promoter (Bartel *et al.*, 1993). By using different promoters, as in PJ69-

4A, this problem can be eliminated. Another common problem is the misfolding of fused proteins, or fused regions blocking interacting domains (Chien *et al.*, 1991). These problems can be overcome by using linker regions or by reversing the fusions on the activating and binding domains. Protein localization and modification is another problem commonly found. However, localization is accomplished in most cases using the nuclear localization signals of the transcriptional activating and binding domains of GAL4.

Results

In order to determine which regions of TraM were involved in dimerization and tetramerization, *traM* was cloned into pGAD-C1 and pGBD-C1 (James *et al*., 1996). This created a fusion between portions of the yeast GAL4 protein, a DNA binding transcriptional activator, and the amino-terminus of TraM. Cloning into pGAD-C1 created a fusion between TraM and the transcriptional activating domain of the GAL4 protein. The full length TraM fused to the activating domain was named pRFAD127 (127 for the full 127 amino acids of TraM). Deletions of the carboxyl-terminus were named according to how many amino acids remained at the amino-terminus (pRFAD119 was a deletion of the carboxyl-terminal 8 amino acids). Deletions of the amino-terminus of TraM were named according to how many amino acids remained on the carboxyl-terminus (pRFADC19 had the final 19 amino acids remaining; C representing carboxyl-terminus). Point mutations in TraM were named by which amino acids were changed (pRFADI109T had the isoleucine at position 109 changed to threonine). Cloning into pGBD-C1 created a fusion between TraM and the DNA binding domain of the GAL4 protein. Nomenclature of the binding domain fusions was identical to the activating domain fusions, however, BD was specified in the name (pRFBD127 vs pRFAD127). The fusions were then placed into PJ69-4A, a *Saccharomyces cerevisiae* strain which contains three reporter genes, the yeast *HIS3* and *ADE2* genes, and the bacterial *lacZ* gene.

Deletional analysis was performed on TraM-activating domain fusion using the *Sal*I and *Eco*RV sites inside the *traM* gene (Figure 5.2). pRFAD119 was cloned from a preexisting TraM clone, pSPE2307 (Penfold, 1995), which had the carboxyl-terminal 8 amino acid deleted. pRFADA37V was cloned from pLRF23 (Penfold, 1995) which had the alanine at position 37 mutated to valine. pRFADI109T was a point mutation changing isoleucine at postion 109 to threonine and was created by a PCR error during cloning. Constructs were only created once, however, both strands of all cloned fragments were sequenced. The presence of the plasmids in the yeast was confirmed by PCR using primers specific for the activating domain and binding domain fusions, followed by restriction enzyme digestion. These clones were tested for their ability to interact with the full length TraM protein fused to the DNA binding region of the GAL4 protein by growth on plates without adenine and histidine. Only the pRFAD127/pRFBD127 (full length AD and full length BD fusions) and pRFADA37V/pRFBD127 combinations grew on these plates. β-galactosidase assays were then performed on all combinations of plasmids and the results are shown in Figure 5.3. Even though β-galactosidase assay results were lower than expected, differences were considered relevant because of the reproducibility of the results. These data show that only the first 24 amino acids (pRFAD24) did not interact with the wild-type TraM BD fusion, all other clones showed some level of interaction in comparison to their controls (with the pGBD-C1 plasmid). Removal of the carboxyl-terminal 19 amino acids (pRFAD108) decreased the strength of the interaction (compare AD127/BD127 to AD108/BD127 which was 13.2 to 1.5 MU) suggesting that this region was important for TraM:TraM interaction. This was confirmed constructing the ADC19/BD127

Figure 5.2. A diagrammatic representation of TraM fusions. The GAL4 transcriptional activating domains are shown in yellow and the DNA binding domains are shown in green. TraM is shown in clear boxes and the *Sal*I and *Eco*RV cleavage sites used in the cloning of the various constructs are shown at the top of the figure. Point mutations are shown as red bars.

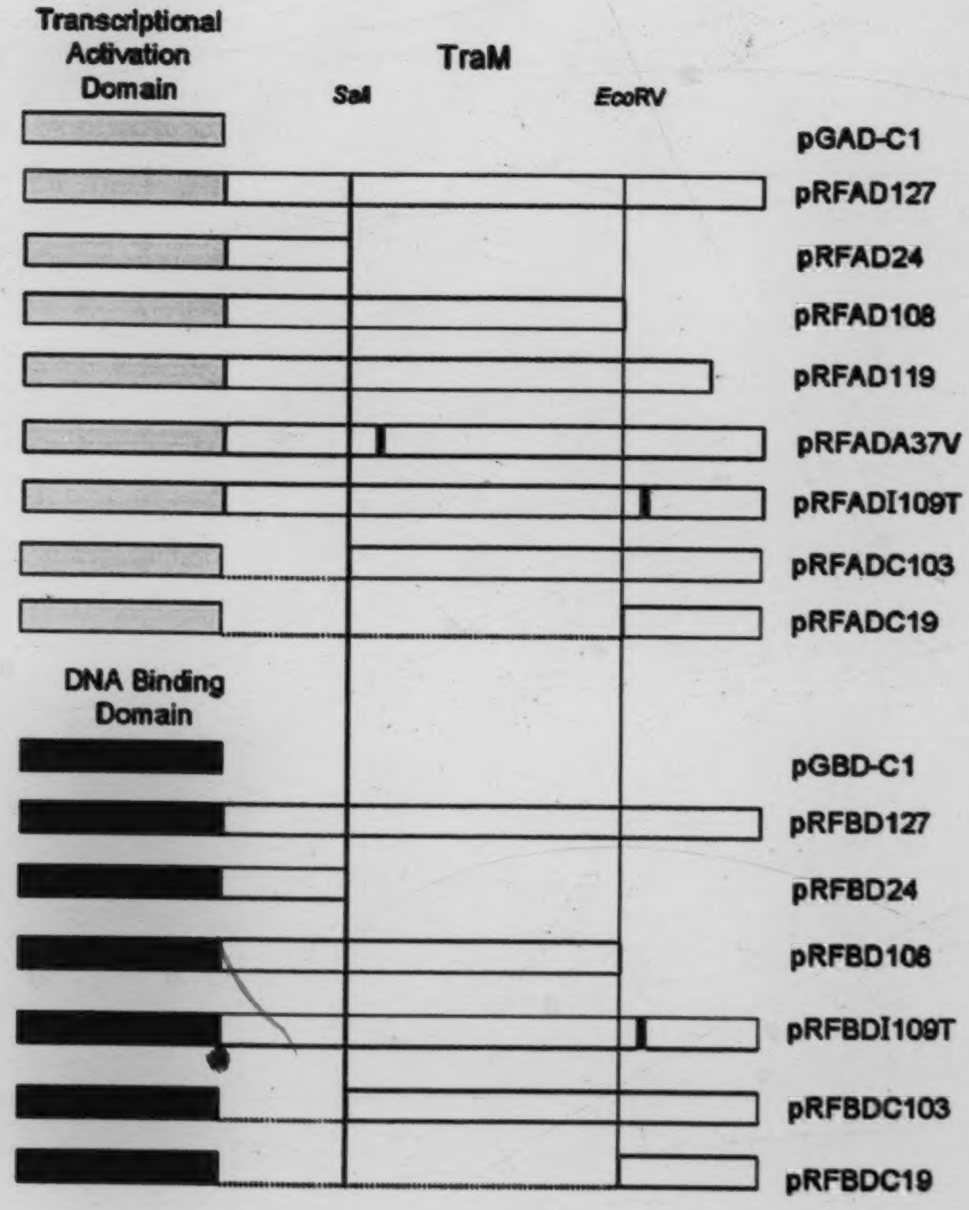
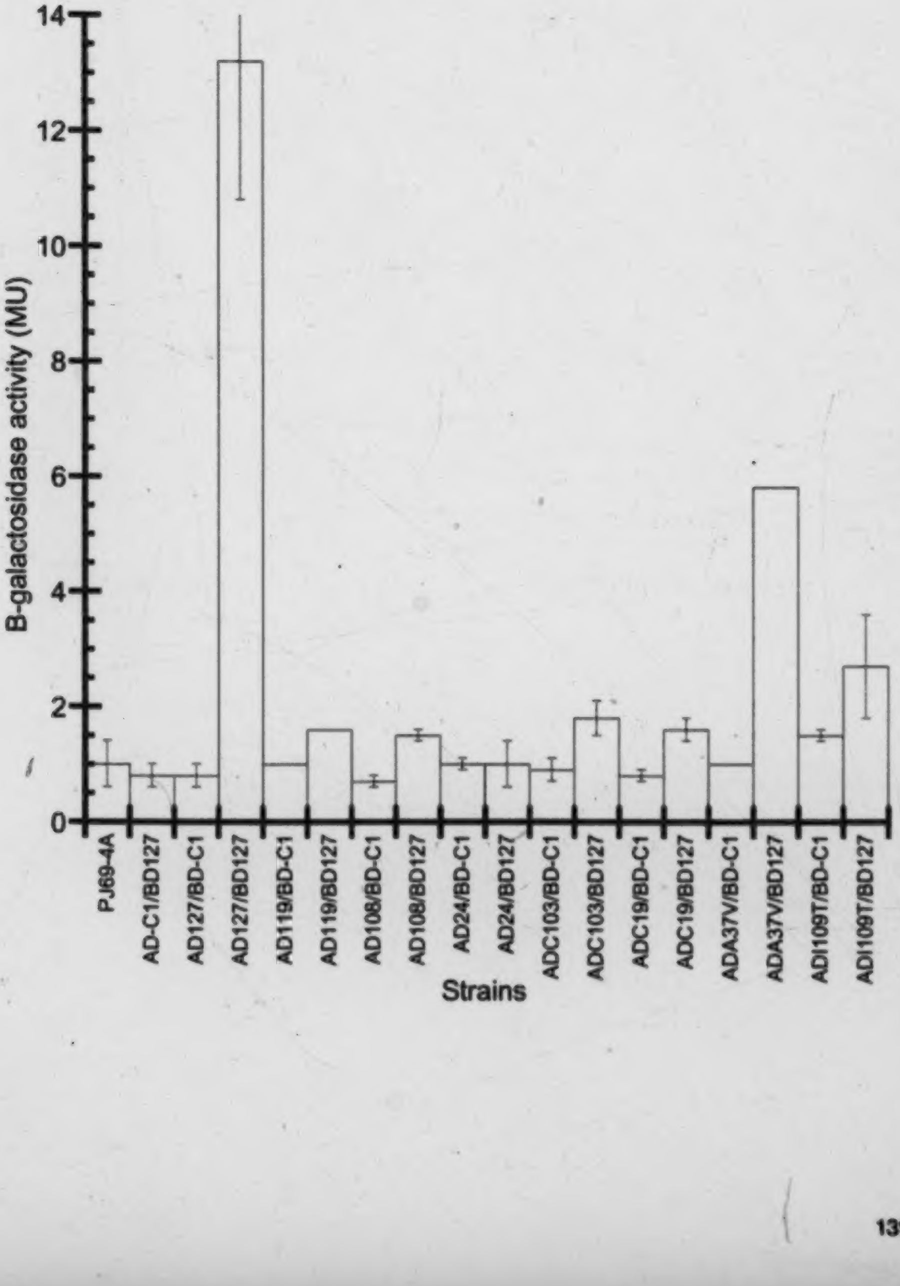

Transcriptional Activation Domain
TraM
SalI
EcoRV
pGAD-C1
pRFAD127
pRFAD24
pRFAD108
pRFAD119
pRFADA37V
pRFADI109T
pRFADC103
pRFADC19
DNA Binding Domain
pGBD-C1
pRFBD127
pRFBD24
pRFBD108
pRFBDI109T
pRFBDC103
pRFBDC19

Figure 5.3. β-galactosidase enzyme activity in extracts of yeast cultures containing the activating domain constructs is shown. Values are means ±SD, and is representative of at least three independent replicates.


B-galactosidase activity (MU)
14
12
10
8
6
4
2
0
PJ69-4A
AD-C1/BD127
AD127/BD-C1
AD127/BD127
AD119/BD-C1
AD119/BD127
AD108/BD-C1
AD108/BD127
AD24/BD-C1
AD24/BD127
ADC103/BD-C1
ADC103/BD127
ADC19/BD-C1
ADC19/BD127
ADA37V/BD-C1
ADA37V/BD127
ADI109T/BD-C1
ADI109T/BD127
Strains

combination which showed some interaction (1.6 MU). Another region of interaction in the central 84 amino acids was suggested when comparing the AD108/BD127 and AD24/BD127 combinations (1.5 and 1.0 MU respectively), and the ADC103/BD127 and ADC19/BD127 combinations (1.8 and 1.6 MU respectively). These two regions are predicted to be α-helical and amphipathic (Figure 5.4) using the Helical Wheel program from GCG. The central α-helix (Figure 5.4a, amino acids 25 to 85) shows a high level of separation of the acidic/basic side and the hydrophobic side. The carboxyl-terminal α-helix (Figure 5.4b, amino acids 101 to 127) has separate hydrophobic and hydrophilic sides, the latter being very acidic. The isoleucine in the I109T clone is boxed in green in Figure 5.4b. Its location in the amphipathic carboxyl-terminal helix, replacing a hydrophobic amino acid, perhaps explains the decreased interaction of pRFADI109T.

In order to determine which region interacted, some of these mutants were cloned into pGBD-C1 which contained the DNA binding region of the GAL4 protein. These clones were then introduced into cells containing the activating domain clones and plated on media lacking adenine and histidine. As before, the AD127/BD127 combinations grew well, however, growth was also observed for other plasmid combinations (Figure 5.5). The pRFAD127/pRFBD108 and pRFAD127/pRFBDI109T combinations grew well, and the pRFAD108/pRFBD108 had moderate growth. Some colonies were also seen with the pRFAD127/pRFBDC103 combination as well. The increased growth is presumably from a higher level of interaction between the fusion proteins. The reason for the increase in interaction of the pRFAD127/pRFBD108 versus the pRFAD108/pRFBD127 combination is presumably due to a conformational problem in the pRFAD108 fusion. This is not

Figure 5.4. Helical wheel analysis of TraM. Hydrophobic amino acids are shown in blue with boxes and non-hydrophobic residues are shown in red. Amino acids 25 to 85 of TraM are represented in A and the alanine at position 37 is indicated with a green box. Amino acids 101 to 127 are represented in B and the isoleucine at position 109 is indicated with a green box.

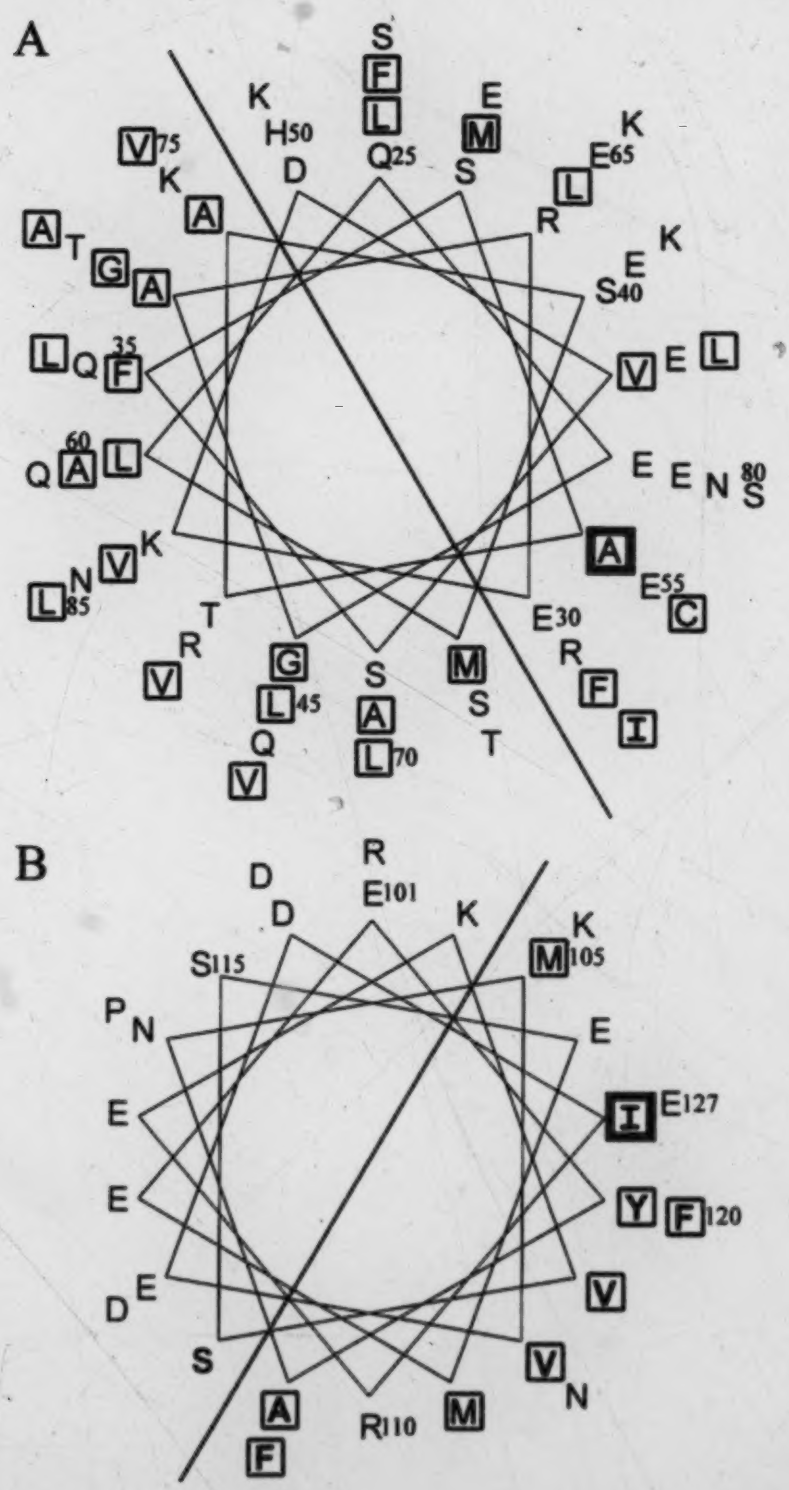
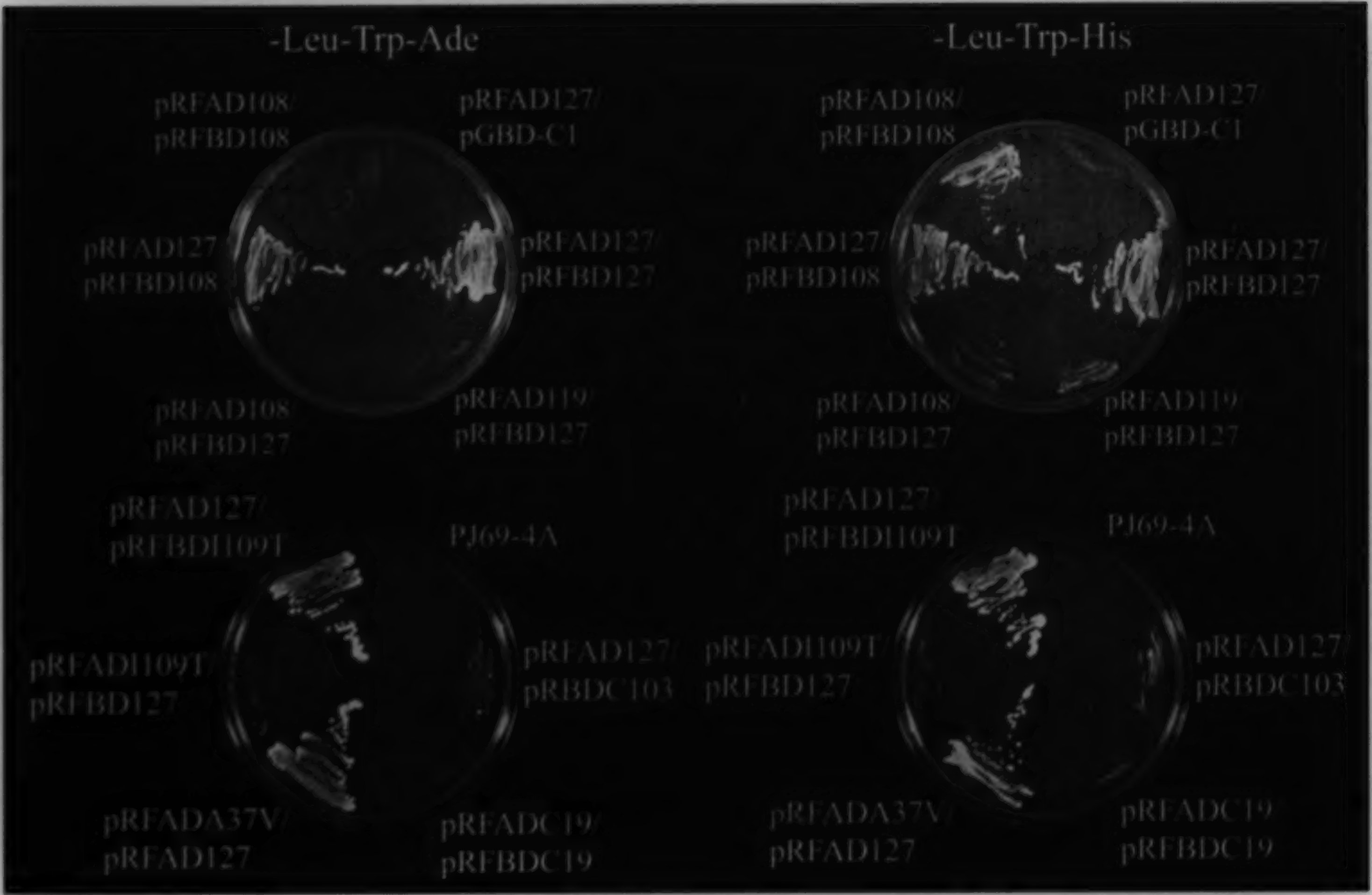

A
S
F
L
Q25
E
M
S
K
H50
D
K
E65
L
R
V75
K
A
A
T
G
A
K
E
S40
L
Q
F
35
V
E
L
60
Q
A
L
E
E
N
80
S
A
E55
C
K
V
N
L
85
T
R
V
E30
R
F
I
G
L
45
Q
V
S
A
L
70
M
S
T
B
R
E101
D
D
K
K
M
105
S115
P
N
E
E
E
I
E127
E
Y
F
120
D
E
V
S
V
N
A
F
R110
M

Figure 5.5. Reporter gene activation as determined by growth of yeast strains on drop-out medium. Medium without leucine, tryptophan and histidine is shown to the right while medium without leucine, tryptophan and adenine is shown to the left. Strains representative of no growth (PJ69-4A), minimal growth (AD108/BD127), moderate growth (AD108/BD108), and good growth (AD127/BD127) are shown.


-Leu-Trp-Ade
pRFAD108/
pRFBD108
pRFAD127/
pGBD-C1
pRFAD127/
pRFBD108
pRFAD127/
pRFBD127
pRFAD108/
pRFBD127
pRFAD119/
pRFBD127
pRFAD127/
pRFBDI109T
PJ69-4A
pRFADI109T/
pRFBD127
pRFAD127/
pRBDC103
pRFADA37V/
pRFAD127
pRFADC19/
pRFBDC19
-Leu-Trp-His
pRFAD108/
pRFBD108
pRFAD127/
pGBD-C1
pRFAD127/
pRFBD108
pRFAD127/
pRFBD127
pRFAD108/
pRFBD127
pRFAD119/
pRFBD127
pRFAD127/
pRFBDI109T
PJ69-4A
pRFADI109T/
pRFBD127
pRFAD127/
pRBDC103
pRFADA37V/
pRFAD127
pRFADC19/
pRFBDC19

uncommon (Clonetech Matchmaker two-hybrid system guide, PT1265-1) and is an example of why most proteins are tested as both AD and BD fusions. β-galactosidase assays were performed on the newly constructed combinations and these results are shown in Figure 5.6. As before, the wild-type TraM fusion, pRFAD127, interacted with all of the binding domain fusions, and weakly with the pRFBD24. As mentioned before, it is clearly noticeable that there is an increased level of interaction of the BD108 and BD109T fusions with the AD127 fusion. The AD108 fusion actually interacts more strongly with the BD108 fusion than with the wild-type fusion (1.8 versus 1.5 MU respectively) suggesting that the central region of TraM interacts with the central region of another TraM molecule. The decrease in β-galactosidase activity when comparing AD108/BDC108 to AD108/BD24 or AD24/BD108 (1.8 to 0.8 or 0.7 MU respectively) suggests that removal of the central region removes the region with which the AD108 fusion interacts. The inability of ADC19 and BD108 fusions to interact (1.1 MU) suggests that the carboxyl-terminus does not interact with the central region of another TraM molecule. The ability of the ADC19/BDC19 fusions to interact (1.7 MU) suggested that the carboxyl-terminus of TraM interacts with the carboxyl-terminus of another TraM molecule. Determination of the fusion protein expression levels was not successful and may be responsible for some of the negative results.

Another way to summarize this data is shown in Figure 5.7 where the data for each combination of clones is averaged. For example the value given for 127/C-1 (column 1) is a combination of all of the AD127/BD-C1 and AD-C1/BD127 data. It can be seen that interactions are the greatest for 127/127 and 108/127 combinations. The reason for the

Figure 5.6. β-galactosidase enzyme activity in extracts of yeast cultures containing binding domain constructs is shown. Values are means ±SD, and is representative of at least three independent experiments.

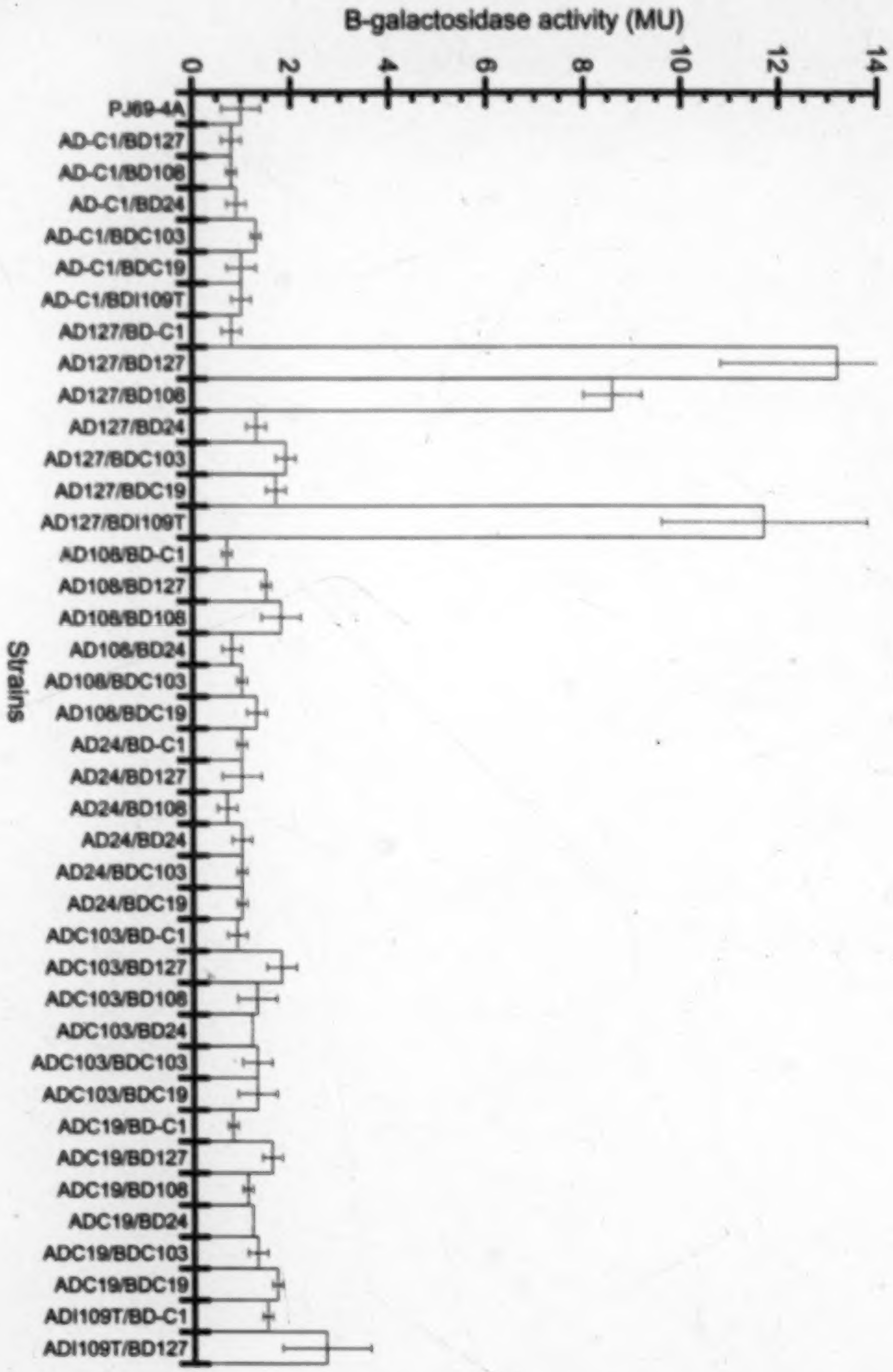
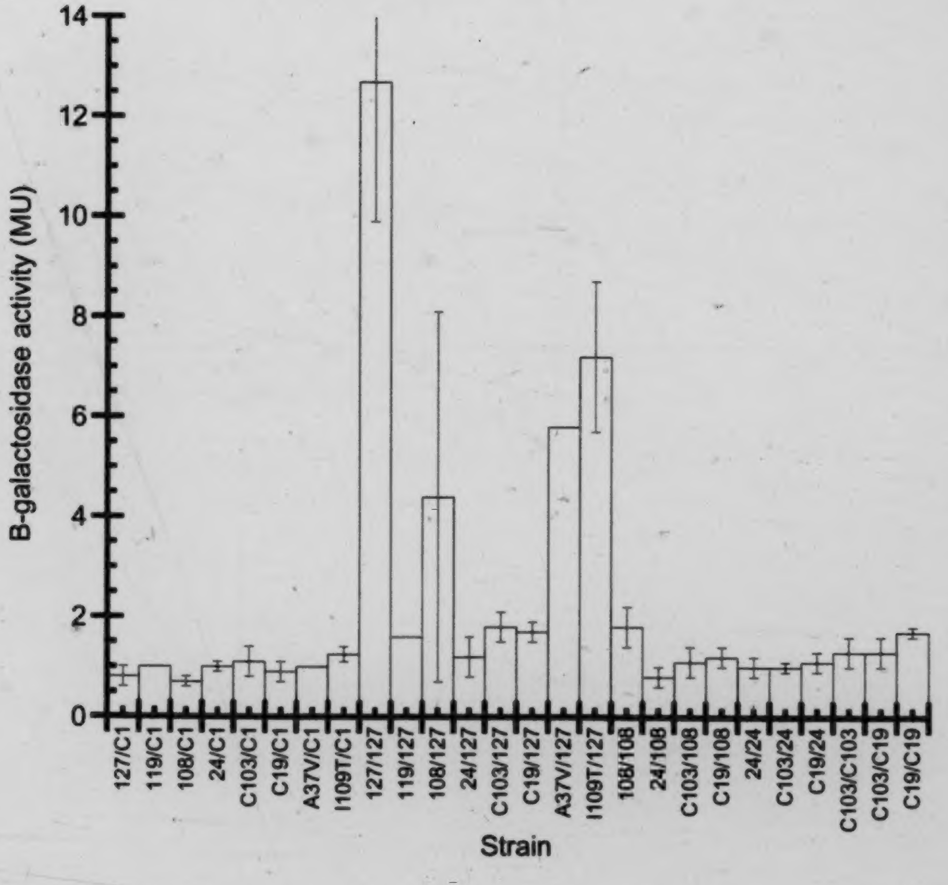

B-galactosidase activity (MU)
0
2
4
6
8
10
12
14
Strains
PJ69-4A
AD-C1/BD127
AD-C1/BD108
AD-C1/BD24
AD-C1/BDC103
AD-C1/BDC19
AD-C1/BDI109T
AD127/BD-C1
AD127/BD127
AD127/BD108
AD127/BD24
AD127/BDC103
AD127/BDC19
AD127/BDI109T
AD108/BD-C1
AD108/BD127
AD108/BD108
AD108/BD24
AD108/BDC103
AD108/BDC19
AD24/BD-C1
AD24/BD127
AD24/BD108
AD24/BD24
AD24/BDC103
AD24/BDC19
ADC103/BD-C1
ADC103/BD127
ADC103/BD108
ADC103/BD24
ADC103/BDC103
ADC103/BDC19
ADC19/BD-C1
ADC19/BD127
ADC19/BD108
ADC19/BD24
ADC19/BDC103
ADC19/BDC19
ADI109T/BD-C1
ADI109T/BD127

Figure 5.7. β-galactosidase enzyme activity in extracts of yeast cultures containing the constructs shown. Results are presented as mixtures of plasmids in each orientration, so data from 108/127 represents all of the data from the AD108/BD127 and AD127/BD108 strains. Values are means ±SD, and is representative of at least three independent replicates.


B-galactosidase activity (MU)
0
2
4
6
8
10
12
14
127/C1
119/C1
108/C1
24/C1
C103/C1
C19/C1
A37V/C1
I109T/C1
127/127
119/127
108/127
24/127
C103/127
C19/127
A37V/127
I109T/127
108/108
24/108
C103/108
C19/108
24/24
C103/24
C19/24
C103/C103
C103/C19
C19/C19
Strain

large error bar for the latter is that the AD108/BD127 combination did not interact while the AD127/BD108 combination did. However, reasons for this are given above. For all other combinations data was reproducible so differences are considered significant. Four other combinations of clones are seen to have a significant level of β-galactosidase activity (hence detectable interaction). These are C103/127, C19/127, 108/108, and C19/C19 combinations (1.8, 1.7, 1.8, and 1.7 MU respectively), all others have less than 1.3 MU of activity. Therefore, as mentioned above, the latter two combinations suggest that the central region of one molecule interacts with the central region of another, and the carboxyl-terminus of one molecule interacts with the carboxyl-terminus of another TraM molecule. These domains can then be suggested to represent the dimerization and tetramerization domains of TraM.

Discussion

The yeast two-hybrid system (Y2HS) has been a valuable tool in the identification of protein:protein interactions *in vivo*, and for the analysis of interacting domains. This system has been used for many types of proteins which cover a wide range of functions. Therefore, this system was utilized to determine which regions of TraM contribute to multimerization.

Analysis using deletions of *traM* fused to the activating domain of *GAL4* identified two regions of interaction for TraM: a central region and a carboxyl-terminal region. The central region (amino acids 25 to 108) was identified by increased β-galactosidase activity of the AD108 and ADC103 clones with the full length TraM in the BD127 clone compared to controls. Subcloning some of the mutations to the binding domain of the GAL4 protein suggested that this region interacts with the same region of another TraM molecule. This is suggested due to the strong growth of the AD108/BD108 combinations on histidine and adenine drop out plates and by the increased β-galactosidase activity of this combination (1.8 MU). This central region is also predicted to be alpha-helical in nature from amino acids 15 to 77 and again from 82 to 89. This region was mapped using the Helical Wheel program (GCG) and appears to be very amphipathic (Figure 5.4A) which suggests that it may function in intra- or intermolecular interactions. The interaction of this central region with the first 24 amino acids was ruled out by the decreased activity of the C103/24 and 108/24 combinations (1.0 and 0.8 MU respectively). This was not surprising since this region of R1 TraM has been designated as the DNA binding region (Verdino *et al.*, 1999). The first 56 amino acids of the R1 TraM protein were shown to form dimers, however, this fragment could only weakly bind to its binding sites (Polzleitner *et al.*, 1997). This region was extended to the first 80 amino acids of R1 TraM using computer generated models and is still predicted to only form dimers (G. Koraimann, personal communication). Since the F and R1 proteins are similar this central region of F TraM may contain the dimerization domain of the protein. The mutation which changed amino acid 37 of TraM from alanine to valine was shown to decrease mating efficiency and DNA binding (Frost *et al.*, 1997) so was also analyzed in this assay. Since this amino acid is located in the amino-terminal portion of the protein, this was interpreted as an affect on the DNA binding portion of the protein. However, β-galactosidase activity of the ADA37V/BD127 combination was decreased to 5.8 MU from 12.8 for the AD127/BD127 combination. Therefore, another interpretation of this mutation may be that the multimerization of the protein is affected, which may also affect DNA binding.

The carboxyl-terminal 19 amino acids of TraM were also suggested by this analysis to contain a multimerization domain. Removal of this region (Figure 5.7, 108/127, 5.4 MU) also greatly decreased β-galactosidase activity compared to the full length TraM molecule (127/127, 12.7 MU). Amino acids 101 to 127 which were predicted to be alpha helical, were plotted using the Helical Wheel program (GCG) and demonstrated a very high degree of amphipathicity (Figure 5.4B). As mentioned above, amphipathic domains are suggested to participate in inter- or intramolecular interactions. Testing this region in the GAL4 binding domain with the activating domain fusions showed that the

ADC19/BDC19 combination interact (Figure 5.7, 1.7 MU). The interaction of the carboxyl-terminal region with the central domain seems unlikely due to the decreased activity of these combinations (Figure 5.7, C19/108 had 1.2 MU). The tetramerization domain of the *E. coli* LacR protein was assigned to the carboxyl-terminal 18 amino acids of the protein. This region is alpha-helical and amphipathic and is attached to the main body of the protein by an extended coil segment of 7 amino acids (Friedman *et al.*, 1995). Tetramerization is thought to occur by the formation of a four-helix bundle from the carboxyl-termini of 4 LacR monomers. The body is also attached to the head-piece which is the DNA binding region and has a helix-turn-helix motif. The dimerization domain of the LacR is located within the main body of the protein and orients the DNA binding domains of both monomers so that they can bind DNA with a higher affinity. The two carboxyl-terminal alpha helices are then oriented in an antiparallel fashion. This is the tetramerization domain and results in the formation of a four-helix bundle containing a leucine heptad repeat which binds together the four monomers (Alberti *et al.*, 1993). Removal of this tetramerization region eliminates LacR's ability to loop DNA and causes an increase in transcription (Brenowitz *et al.*, 1991; Oehler *et al.*, 1990). The carboxyl-terminal region of R1 TraM also appears to follow this type of arrangement with two carboxyl-terminal alpha helices linked to the body of the protein by a 19 amino acid loop linker (Verdino *et al.*, 1999). Using computer modeling these alpha helical regions also seem to arrange into a four-helix bundle in a tetramer of R1 TraM (G. Koraimann, personal communication). Similarly F TraM has an extended region of approximately eleven amino acids preceding the carboxyl-terminal alpha helix. The amphipathic carboxyl-terminal alpha helix and similarity to LacR and R1 TraM suggests that the carboxyl-terminal alpha helix may be the tetramerization domain of the F TraM molecule. The removal of the carboxyl-terminal eight amino acids of the F TraM protein resulted in a nonfunctional protein, unable to complement a F TraM mutant plasmid (Penfold, 1995). The ability of the carboxyl-terminal eight amino acid deletion to interact with the wild type TraM (AD119/BD127, 1.6 MU, Figure 5.5) in a similar manner as the larger deletion (AD108/BD127, 1.5 MU, Figure 5.5) suggests that removal of even the last eight amino acids destroys the carboxyl-terminal region of protein:protein interaction. The inablility of the carboxyl-terminal eight amino acid deletion of TraM to function *in vivo* may be that the tetramerization domain of TraM has been altered, and is now unable to perform it function.

Chapter 6

Characterizing the DNA binding activity of TraM

Introduction

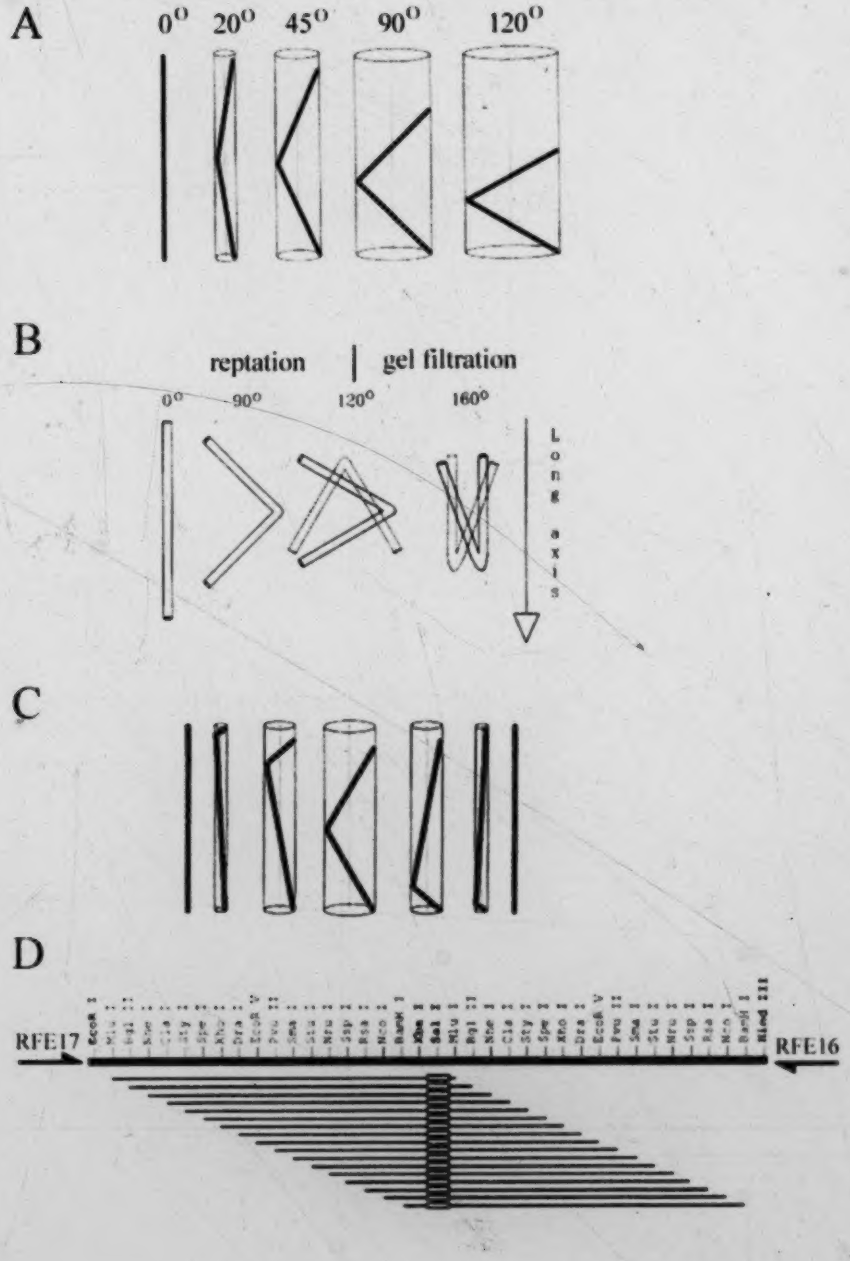
Using a purified preparation of TraM, the binding sites in the *oriT* region were defined using DNase I footprinting (Di Laurenzio *et al*., 1992). Strong footprints were found on the top (non-transferred) strand suggesting that TraM bound only the upper strand. TraM binding to the entire *oriT* region was calculated to have a K_a of approximately 10^7 M^{-1} (Di Laurenzio *et al*., 1992) and EMSA experiments showed that purified TraM gave only 2 different species when binding *oriT* (Penfold, 1995). This was unexpected since *oriT* contains three TraM binding sites. It was proposed that the first species represents a tetramer of TraM bound to *sbmA* and the second species represents a second tetramer binding to *sbmB*. Complete retardation was seen at approximately 2 nM of purified protein. Crude cell extracts containing TraM were also used for analysis of the TraM binding sites *sbmC* and *sbmBC*. *sbmC* alone had a much lower affinity for TraM and bands tended to smear. This suggested that the protein:DNA interactions were continuously dissociating and reforming. During EMSA, TraM had a higher affinity for *sbmBC* and distinct bands were observed. This suggested that binding of TraM to *sbmBC* was more stable than to *sbmC*. However, the affinity of TraM for *sbmBC* was still much lower than for a fragment containing all three binding sites. Many of the experiments were not comparable since crude and purified extracts containing TraM had been used. To understand the mechanism by which TraM binds to its binding sites, a more thorough characterization of the binding was required. This requires the analysis of the three TraM binding sites alone and in various combinations using purified protein preparations.

Comparing this binding data can lead to a model for how TraM binds to *oriT*. This type of thorough study has not been performed, and is required for a complete analysis.

Electrophoretic mobility shift assays (EMSA) are a useful tool for the analysis of interactions between proteins and DNA (Carey, 1991). An advantage over other binding assays, such as filter binding assays, is that during EMSA, different bound complexes can be separated from each other. These can suggest a mode for protein binding. Many factors affect the running of DNA such as molecular weight, net charge, conformation of the bound DNA, composition and concentration of gel matrix, and temperature (Lane *et al.*, 1992). Complicating the analysis of complexes is that more than one of these factors can vary as more protein is bound to the DNA. Molecular weight of the retarded species can be calculated if DNA conformation is not affected by binding. In these cases addition of protein can have a predicted effect on the relative mobility of the complex. Truncated or fusion proteins can also decrease or increase relative mobility of fragments, respectively, and assist in the calculation of molecular weights. Larger DNA fragments can decrease relative mobility, leading to the conclusion that it is ratio of the protein to DNA masses rather than the final mass of the complex that determines relative mobility (Lane *et al.*, 1992). Changes in protein charge can also affect mobility of the protein:DNA complexes. Experiments were done with the Trp repressor where the pH was varied and conditions above the pI of the protein caused an increased mobility of the retarded species (Carey, 1988). Conformation of the bound DNA also affects mobility. For example, bending of the DNA can cause a decrease in mobility of the complex. Mobility depends on the bend angle and on the position of the bend in the DNA fragment. Increases in gel concentration can also decrease relative mobility since larger molecules are more affected by increases in gel concentration. Examples of EMSA involve many types of DNA binding proteins such as transcriptional repressors like LacR (Brenowitz *et al.*, 1991), replication initiator proteins like DnaA (Weigel *et al.*, 1997) and restriction enzymes like *EcoR*V (Taylor *et al.*, 1991).

Curved DNA structures can cause a decrease in the migration of DNA (Koo and Crothers, 1988). Mobility change is actually determined by the average distance between the DNA ends, so that a decrease in distance can cause a decrease in the mobility of the DNA (Lane *et al.*, 1992). This can be demonstrated by using a cylinder model where the decrease in end to end distance actually increases the cylinder's diameter (Figure 6.1a). The larger cylinder diameter then causes an increase in the congestion in the gel pores. This effect reaches a maximum when the bend angle is 120^{O}. Bend angles above this are hypothesized to cause a reorientation of the molecule during movement through the gel so that separation becomes based on gel filtration and not by reptation of the DNA (Zinkel and Crothers, 1990), as shown in Figure 6.1b. As mentioned above, the effects of the bend angles are also increased as gel concentration is increased (Marini *et al.*, 1982). Another phenomenon is also seen by changing the position of the bend in the DNA fragment. As shown in Figure 6.1c, as the bend is moved from the center of the fragment to the end, the cross-sectional area of the molecule is decreased, causing an increase in the mobility of the DNA (Lane *et al.*, 1992). This last phenomenon is the basis for the calculation of bend angle using the plasmid pBEND2 (Kim *et al.*, 1989). In these assays a DNA binding site is cloned into the center of the multiple cloning site of the plasmid.

Figure 6.1. A. The bend angle effects end to end distance of DNA molecules. Cylinders are shown around each molecule to show the increase in the cross-sectional diameter as bend angle increases. B. The change in long axis changes as bend angle increases beyond 120^{o}. Up to 120^{o} the long axis (direction of movement of the molecule) is parallel to the line connecting the ends of the molecule. Beyond 120^{o} the long axis changes to become parallel to one of the arms of the molecule. C. End to end distance increases as the position of the bend moves from the center to one end. As the distance increases the cross-sectional area of the molecule decreases, increasing the mobility of the molecule during electrophoresis. D. Position of restriction enzyme sites in the multiple cloning site of pBEND2. The *Sal*I and *Xba*I sites can be used for cloning binding sites (shown in a box). The duplicated sites can be used to produce equal length fragments with the cloned sequences at various positions in the fragments shown below. Position of primers RFE16 and RFE17 which were used in PCR reactions to amplify the binding site clones are shown. Figures A, B and C are adapted from Lane *et al.*, 1992, and D is adapted from Kim *et al.*, 1989.


A
0°
20°
45°
90°
120°
B
reptation
gel filtration
0°
90°
120°
160°
Long axis
C
D
RFE17
RFE16
EcoR I
Mlu I
Bgl II
Xho I
Cla I
Sty I
Spe I
Xba I
Dra I
EcoR V
Pvu II
Sma I
Stu I
Nru I
Ssp I
Rsa I
Nco I
BamH I
Xba I
Sal I
Mlu I
Bgl II
Nhe I
Cla I
Sty I
Spe I
Xho I
Dra I
EcoR V
Pvu II
Sma I
Stu I
Nru I
Ssp I
Rsa I
Nco I
BamH I
Hind III

When cut with a variety of enzymes (each giving DNA fragments of equal length) the cloned sequence is located in an assortment of positions ranging from one end to the other. By comparing the mobility of the retarded species having the binding site at the end to one having the binding site in the middle can give the angle by which the DNA is bent by the bound protein. This technique can also be used to find intrinsic bends in the DNA (Koo and Crothers, 1988).

Identification of protein binding sites on DNA has been greatly assisted by a method known as footprinting. Footprinting uses chemical reagents or enzymes to identify which nucleotides are bound by a protein. A variety of footprinting techniques have become available, however, a method known as hydroxyl radical footprinting has become quite useful (Tullius *et al.*, 1987). This technique uses highly reactive hydroxyl radicals (•OH) to remove deoxyribose hydrogens from the DNA backbone. This results in the cleavage of the backbone and after separation on denaturing gels, cleavage sites can be correlated to specific bases. One of the greatest advantages of this procedure is that the hydroxyl radicals show no base or sequence specificity (Tullius and Dombroski, 1985; Tullius and Dombroski, 1986), however, the exact mechanism for cleavage is still unknown. The hydroxyl radicals are produced using the Fenton reaction which uses iron(II)EDTA to reduce hydrogen peroxide and produces hydroxyl radicals, hydroxide ions, and iron(III)EDTA. The iron(III)EDTA is then reduced back to iron(II)EDTA using ascorbate, allowing a cyclic reaction to be maintained. Hydroxyl radicals are also produced from ionizing radiation and can result in DNA damage. It has been shown that solvent accessibility is one of the factors that determines which deoxyribose hydrogens are reactive, and that these are most likely the hydrogens on the 5' and 4' carbons on the DNA backbone (Balasubramanian *et al.*, 1998). Hydroxyl radicals have been used for footprinting protein:DNA complexes (Dixon *et al.*, 1991), and DNA and RNA secondary structure (Price and Tullius, 1992; Latham and Cech, 1989, respectively). Signal intensity can be problematic, however, and can be increased by isolating bound DNA species after hydroxyl radical reactions using EMSA (Dixon *et al.*, 1991). This decreases background levels by removing the unbound DNA which has been cleaved. Other footprinting reagents include dimethyl sulfate which methylates guanine and adenine residues and DNase I footprinting which uses an enzyme to cleave the DNA, and so does not give as fine a footprint as smaller cleavage molecules like hydroxyl radicals. Methidium propyl-EDTA-iron(II) footprinting also uses the Fenton reaction (Hertzberg and Dervan, 1984; Tullius *et al.*, 1987). The methidium intercalates into the DNA bringing the iron(II) close to the DNA backbone which then reacts with hydrogen peroxide to produce hydroxyl radicals which cleave the DNA. This technique requires the intercalation of methidium which does not occur where protein is bound. This results in broader footprints much like those seen with DNase I. Another method uses bis(1,10-phenanthroline)copper(I) which binds to the DNA and allows the copper to cleave the DNA in a method similar to the Fenton reaction (Spassky and Sigman, 1985; Tullius *et al.*, 1987). This method also produces broad footprints like those seen with DNase I, presumably due to its requirement to bind to DNA before cleavage.

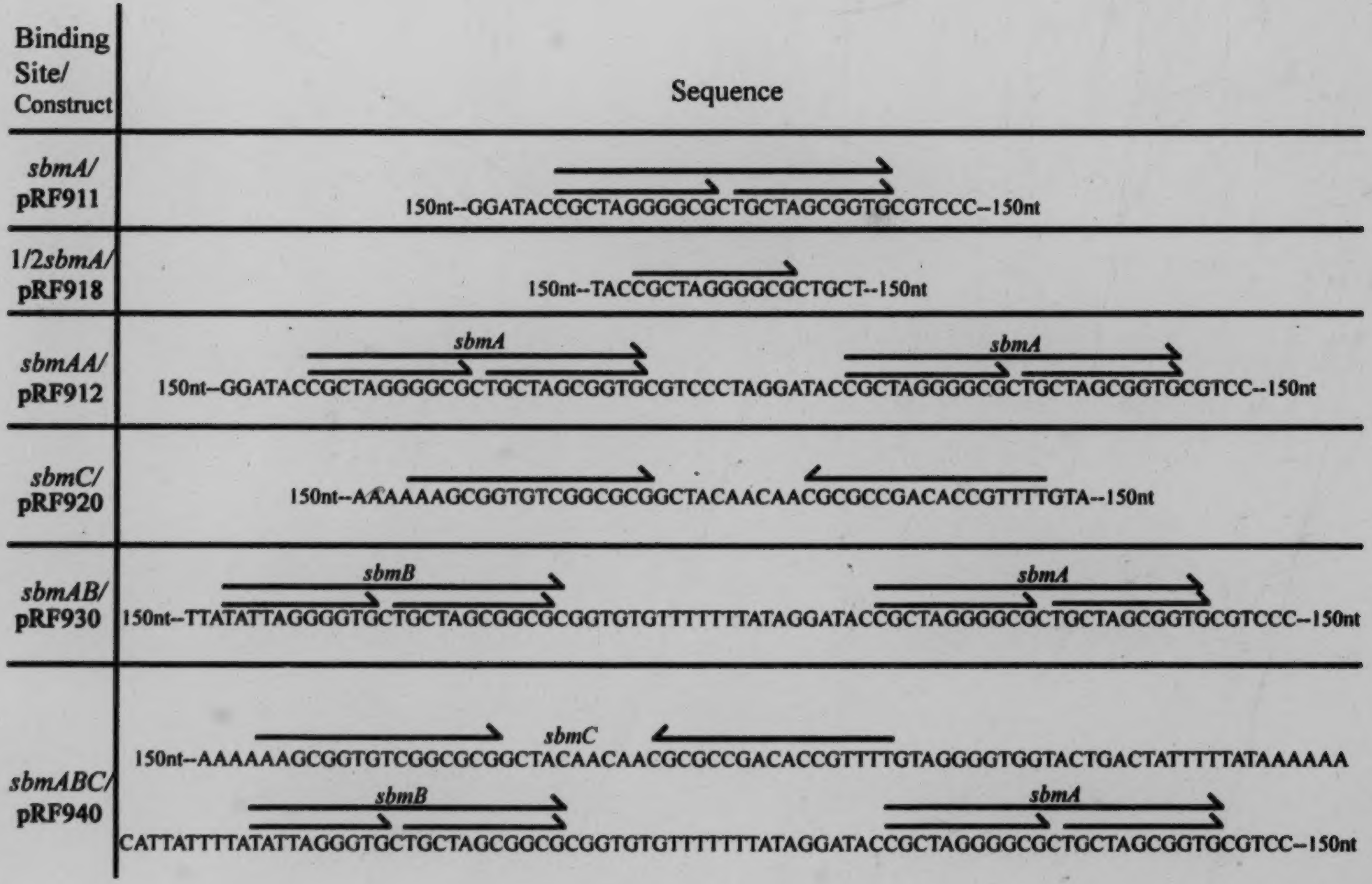
Results

Electrophoretic mobility shift assays.

The TraM binding sites were cloned individually and in combinations into pBEND2 (Kim *et al.*, 1989) to give pRF911 (*sbmA*), pRF912 (tandem repeat of *sbmA*), pRF918 (half of the *sbmA* site), pRF920 (*sbmC*), pRF930 (*sbmAB*), and pRF940 (*sbmABC*). The sequences for the cloned sites are given in Table 6.1. DNA used in EMSA was produced by PCR using radioactively end-labeled primers annealing inside the pBEND2 multiple cloning site, or by using cold primers with [α^{32}P]dCTP in the reaction. This DNA was purified from agarose gels, and was quantified using a scintillation counter. Approximately 0.3 fmol of DNA (approximately 3000 cpm) was used in each binding reaction. Binding reactions contained DNA, the appropriate amount of purified TraM in TED buffer, binding buffer, and 1 μg of non-specific competitor DNA (poly dIdC). Reactions were incubated at 37^{O}C for 15 minutes and then loaded onto an 8% polyacrylamide gel which had been pre-run at 4^{O}C for 15 minutes at 25 mA. Gels were run at 25 mA at 4^{O}C until the bromophenol blue dye reached the bottom of the gel, they were dried, and placed on a Molecular Dynamics Phosphor screen overnight. The screen was analyzed using a Molecular Dynamics Phosphor Imager 445SI and bands were quantitated using Image QuaNT v.4.2a software.

Examples of EMSA using all of the binding site combinations and negative control are shown in Figures 6.2 and 6.3. EMSA were also performed using finer dilutions of TraM and are shown in Figure 6.4 and 6.5. It can be seen that TraM binds to *sbmA*, *sbmAB*,

Table 6.1. Sequences of the TraM binding sites cloned into pBEND2 to be used for EMSA and bending experiments. The positions of putative repeats are also shown with red and blue arrows. On larger cloned sites such as *sbmAA*, *sbmAB*, and *sbmABC* the positions of the individual binding sites are indicated. The number of nucleotides out to RFE16 and RFE17 are also shown.



Binding Site/ Construct	Sequence
sbmA/ **pRF911**	150nt--GGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCCC--150nt
1/2*sbmA*/ **pRF918**	150nt--TACCGCTAGGGGCGCTGCT--150nt
sbmAA/ **pRF912**	150nt--GGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCCCTAGGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCC--150nt
sbmC/ **pRF920**	150nt--AAAAAAGCGGTGTCGGCGCGGCTACAACAACGCGCCGACACCGTTTTGTA--150nt
sbmAB/ **pRF930**	150nt--TTATATTAGGGGTGCTGCTAGCGGCGCGGTGTGTTTTTTTATAGGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCCC--150nt
sbmABC/ **pRF940**	150nt--AAAAAAGCGGTGTCGGCGCGGCTACAACAACGCGCCGACACCGTTTTGTAGGGGTGGTACTGACTATTTTATAAAAAA CATTATTTTATATTAGGGTGCTGCTAGCGGCGCGGTGTGTTTTTTTATAGGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCC--150nt

Figure 6.2. Binding of TraM to *sbmA*, *sbmC*, *sbmAB*, and *sbmABC* to determine relative affinities. Binding sites were cloned into pBEND2 and bound by the indicated amounts of TraM shown in nM. Each gel was run individually and band positions are not comparable. Each lane contains 0.3 fmol of DNA which was produced by PCR using end-labeled primers.

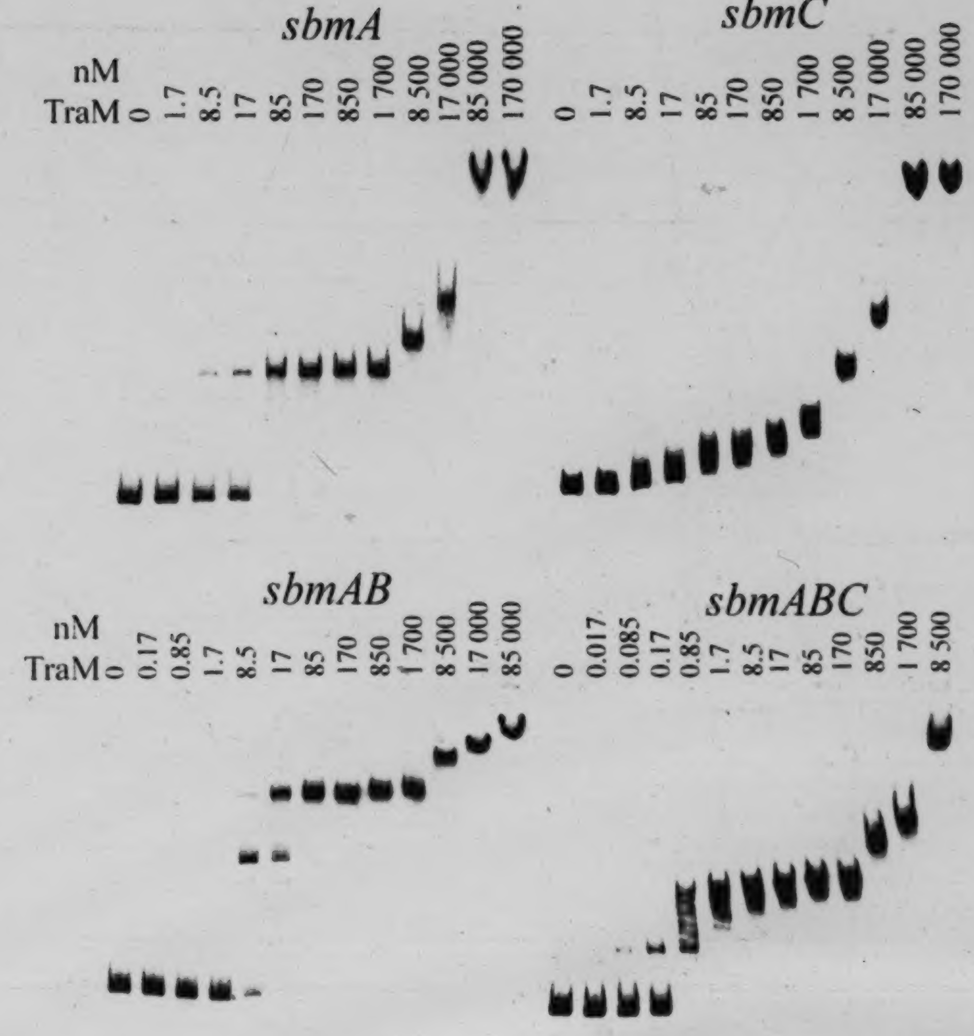
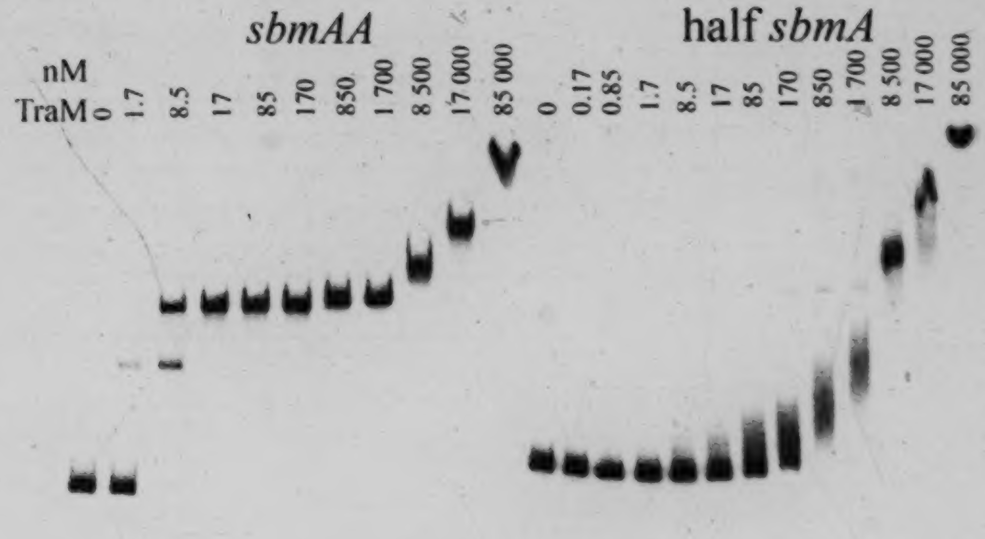

sbmA
nM
TraM
0
1.7
8.5
17
85
170
850
1 700
8 500
17 000
85 000
170 000
sbmC
0
1.7
8.5
17
85
170
850
1 700
8 500
17 000
85 000
170 000
sbmAB
nM
TraM
0
0.17
0.85
1.7
8.5
17
85
170
850
1 700
8 500
17 000
85 000
sbmABC
0
0.017
0.085
0.17
0.85
1.7
8.5
17
85
170
850
1 700
8 500

Figure 6.3. Binding of TraM to *sbmAA*, half *sbmA*, and a negative control to determine relative affinities. Binding sites were cloned into pBEND2 and bound by the indicated amounts of TraM shown in nM. The negative control is the vector only, without cloned DNA. Each gel was run individually so band positions should not be compared. Each lane contains 0.3 fmol of DNA which was produced by PCR using end-labeled primers.


sbmAA
half sbmA
nM
TraM
0
1.7
8.5
17
85
170
850
1 700
8 500
17 000
85 000
0
0.17
0.85
1.7
8.5
17
85
170
850
1 700
8 500
17 000
85 000

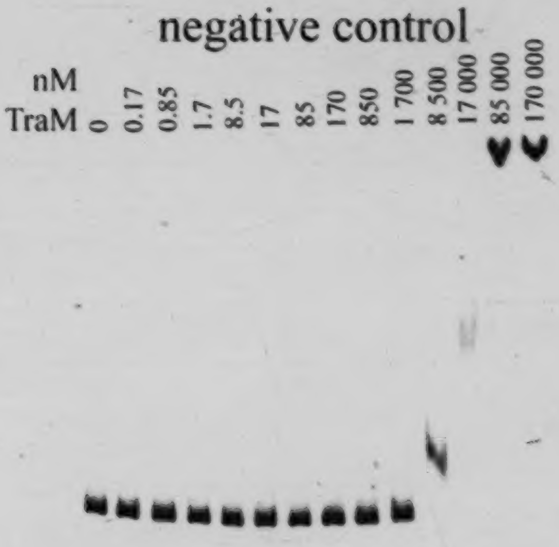

negative control
nM
TraM
0
0.17
0.85
1.7
8.5
17
85
170
850
1 700
8 500
17 000
85 000
170 000

Figure 6.4. Binding of TraM to *sbmA*, negative control, and *sbmC* cloned into BEND2 using a finer range of TraM dilutions (shown in nM). Each lane contains 0.3 fmol of DNA which was produced by PCR using end-labeled primers.

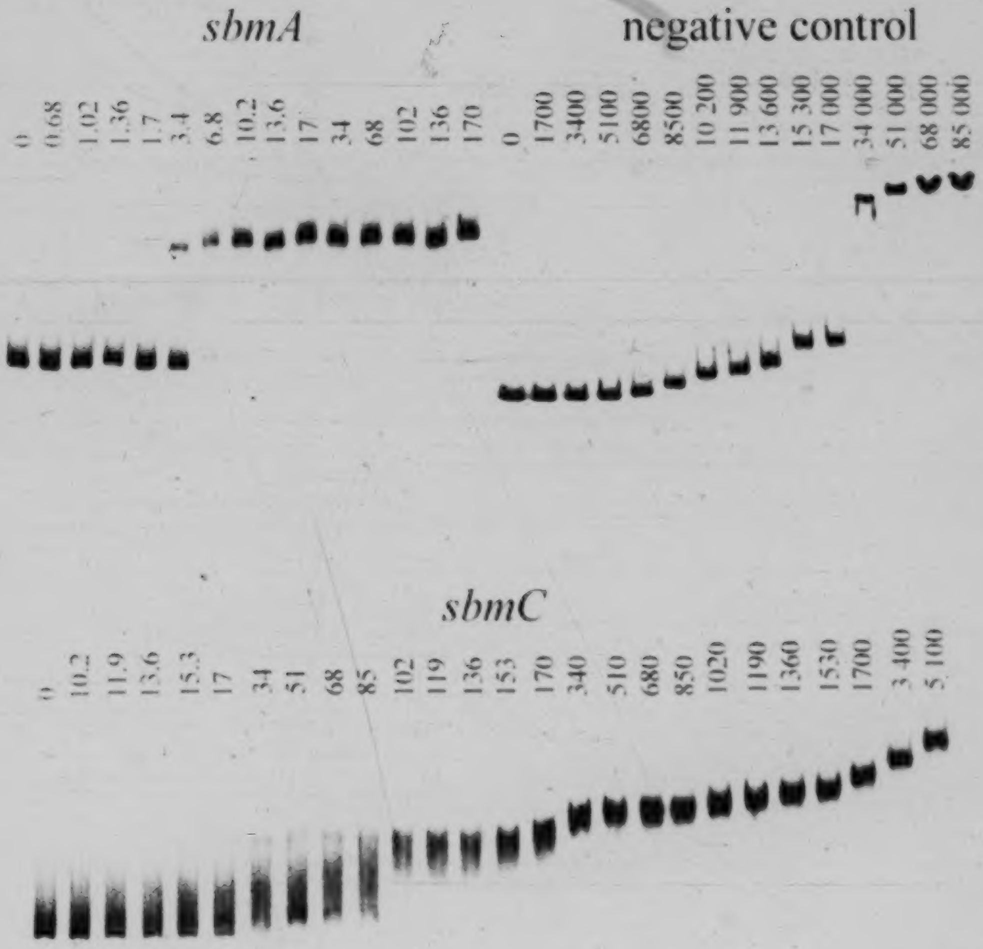
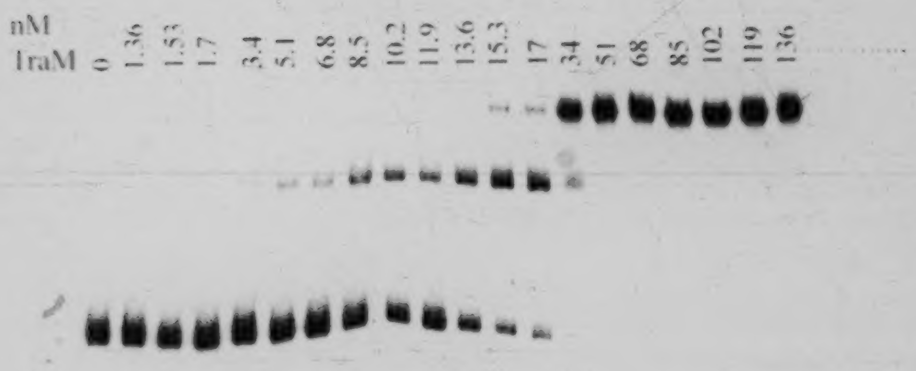
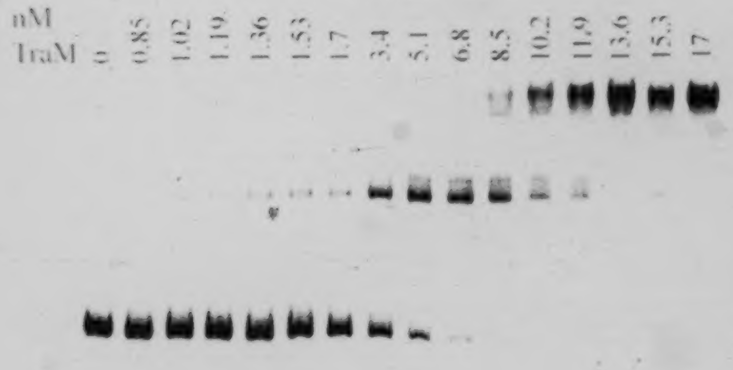

sbmA
0
0.68
1.02
1.36
1.7
3.4
6.8
10.2
13.6
17
34
68
102
136
170
negative control
0
1700
3400
5100
6800
8500
10 200
11 900
13 600
15 300
17 000
34 000
51 000
68 000
85 000
sbmC
0
10.2
11.9
13.6
15.3
17
34
51
68
85
102
119
136
153
170
340
510
680
850
1020
1190
1360
1530
1700
3 400
5 100

Figure 6.5. Binding of TraM to *sbmAB* and *sbmABC* cloned into BEND2 using a finer range of TraM dilutions (shown in nM). Each lane contains 0.3 fmol of DNA which was produced by PCR using end-labeled primers.



sbmAB

nM
TraM
0
1.36
1.53
1.7
3.4
5.1
6.8
8.5
10.2
11.9
13.6
15.3
17
34
51
68
85
102
119
136



sbmABC

nM
TraM
0
0.85
1.02
1.19
1.36
1.53
1.7
3.4
5.1
6.8
8.5
10.2
11.9
13.6
15.3
17

and *sbmABC* in an ordered fashion giving defined species during EMSA. However, the size of the protein multimer bound to *sbmC* only increases in size as more TraM is added to the reactions. The unbound DNA in each lane was quantitated for each DNA fragment and divided by the value obtained from the lane containing no TraM. This gave the percentage of unbound DNA, which is graphed versus TraM concentration in Figure 6.6A. The association constants (K_a) were calculated from the binding curves (1/[TraM] at 50% unbound) and are shown in Table 6.2. This data shows that TraM binds *sbmA* and *sbmAB* with similar affinity, and *sbmC* with weaker affinity. TraM binds 5 times more strongly to *sbmABC* than to *sbmAB* and almost 10 times more strongly than to *sbmC*. The fact that affinity is increased as the DNA fragment contains more binding sites suggests that TraM binds all three sites together in a cooperative fashion. Interestingly TraM binds *sbmAA* (a direct repeat of *sbmA*, and hence similar to *sbmAB*) with similar affinity as *sbmA* and *sbmAB* suggesting that *sbmAB* is similar to two *sbmA* sites in a direct repeat orientation. Quantification of fully bound complexes of *sbmA*, *sbmAB*, *sbmABC*, and *sbmAA* were performed and graphed (Figure 6.6B). It can be seen that all four binding site combinations form fully bound complexes (Figure 6.6B) with the same profile that they bind free DNA (Figure 6.6A; sbmAB slightly slower than the other three). However, higher protein concentrations are required to obtain 50% binding (approximately 2 to 3-fold). For binding sites which gave more than one species during EMSA (*sbmAB*, *sbmAA*, and *sbmABC*) graphs were also constructed with lines to represent each species. Quantitated values for bands in each lane were totaled to give a lane total. The value for each band was then divided by this lane total. This way the value for each band represents a percentage of the total DNA in the lane. The graph for

Figure 6.6. Graphs demonstrating the affinity of TraM for its binding sites. Figure legends designate the binding sites by colour; *sbmA*: dark blue; *sbmAA*: red; half *sbmA*: green; *sbmAB*: grey; *sbmABC*: light blue; negative control: orange; single-stranded *sbmA* upper strand: brown; *sbmC*: black. A. Data is presented as the percent of DNA remaining in an unbound state at each protein concentration. Y-axis represents the amount of unbound DNA remaining at each protein concentration (X-axis) and is shown in nM. Error bars represent the standard deviation of data from multiple gels. B. Data showing the accumulation of the fully bound complex at each protein concentration (shown in nM) and is shown as a percentage of the total DNA in each lane. Buildup varies based on affinity and on the accumulation of intermediate complexes.

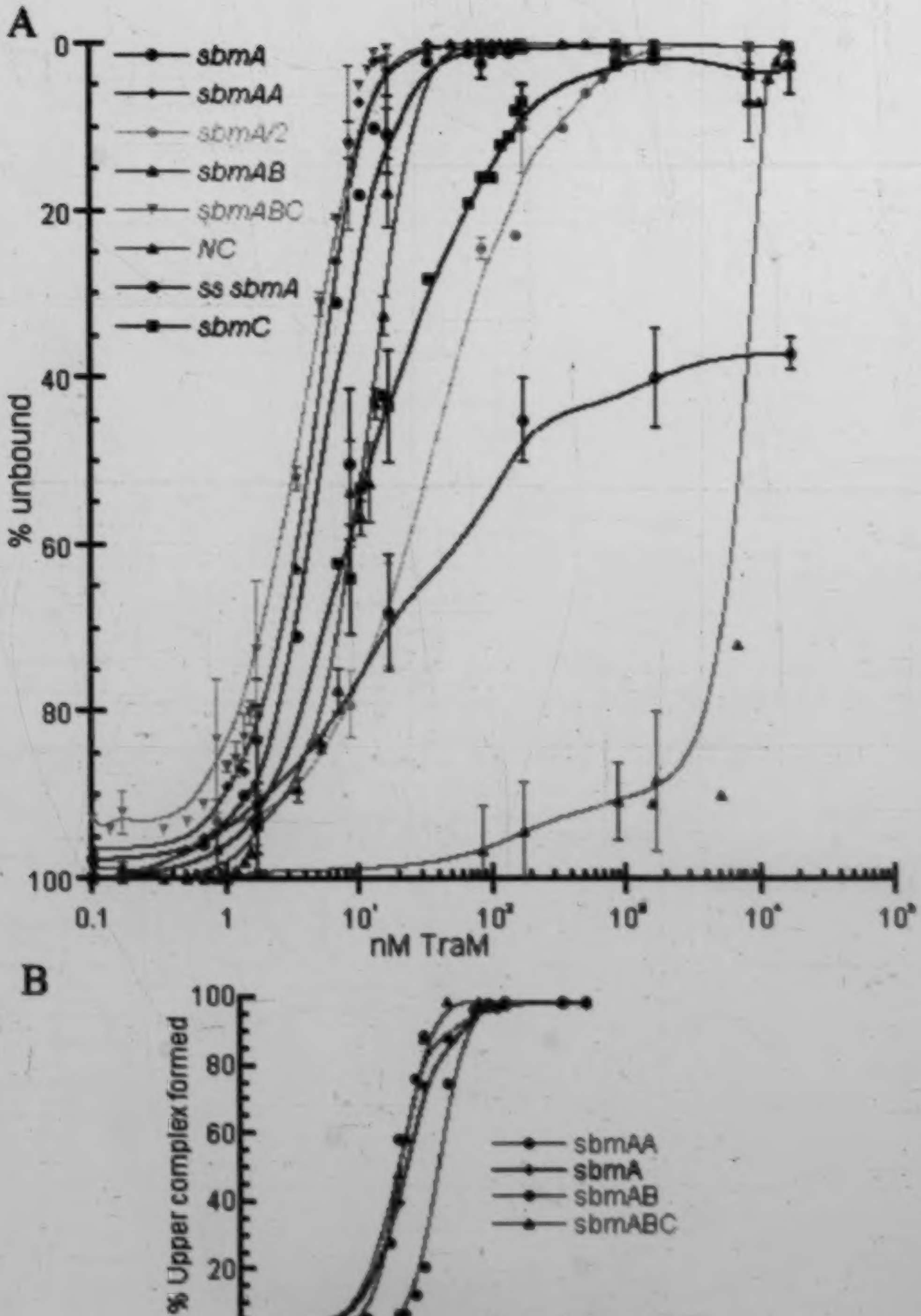

A
sbmA
sbmAA
sbmA/2
sbmAB
sbmABC
NC
ss sbmA
sbmC
% unbound
0
20
40
60
80
100
0.1
1
10¹
10²
10³
10⁴
10⁵
nM TraM
B
% Upper complex formed
100
80
60
40
20
0
0.1
1
10
100
1000
10000
nM TraM
sbmAA
sbmA
sbmAB
sbmABC

Table 6.2

Quantitative analysis of TraM binding to the TraM binding sites

Binding Site	K_a (M^{-1})[a]	Breadth of binding curve[b]	τ[b]	ω[b]
sbmA	2×10^8	1.27	n/d	169
sbmC	6.7×10^7	1.94	n/d	n/d
sbmAB	1×10^8	0.98	1.3±0.3	4
sbmABC	5×10^8	1.04	4.4±2.3	n/d
1/2sbmA	1.5×10^7	2.1	n/d	n/d
sbmAA	2.5×10^8	1.07	2.3±1.0	1.6

[a] value represents formation of all complexes

[b] discussed in section on cooperativity

sbmAB is shown in Figure 6.7 and demonstrates the buildup of the middle band. However, this middle species never comprises all the DNA present in the reaction. This suggests that an initial intermediate species is formed during binding, but that a second final species is then quickly formed.

DNaseI footprinting experiments suggested that TraM bound to only one strand (DiLaurenzio *et al*., 1992). Therefore, the ability to bind single-stranded DNA during EMSA was assayed. Oligonucleotides (summarized in Table 6.3) representing the entire upper strand (RFE12) and lower strand (RFE11) of *sbmA* (non-transferred and transferred strands respectively) were synthesized, radioactively end-labeled, and used in EMSA. Figure 6.8 shows that the upper strand, but not the lower strand, is bound by TraM. The appearance of an intermediate band (arrow) in the single-stranded EMSA reactions suggested that RFE12 (the upper strand) was either dimerizing or forming some type of secondary structure. The disappearance of this species when TraM was added raised the possibility that it was this species that TraM was binding and not to single-stranded DNA. Quantitation of the unbound species was done and is plotted in Figure 6.6A. It can be seen that even though TraM binds to the upper strand, it does so at a reduced affinity with an association constant of approximately $6.7 \times 10^6 M^{-1}$. Duplexes of the upper and lower strand were made (Figure 6.9), and bound by TraM with a similar affinity as the upper strand alone.

To determine if the intermediate species was formed from intra-molecular interactions, various DNA folding programs such as DNA *mfold* were used (Santa Lucia, 1998) to

Figure 6.7. Graph representing the buildup and loss of the various *sbmAB* complexes. Data shows unbound (red), middle band or intermediate complex (blue), and upper complexes (green) at different TraM concentrations (shown in nM). Quantification of bands was performed and is shown as a percentage of the total DNA in each lane. Error bars represent the standard deviation of data from multiple gels.

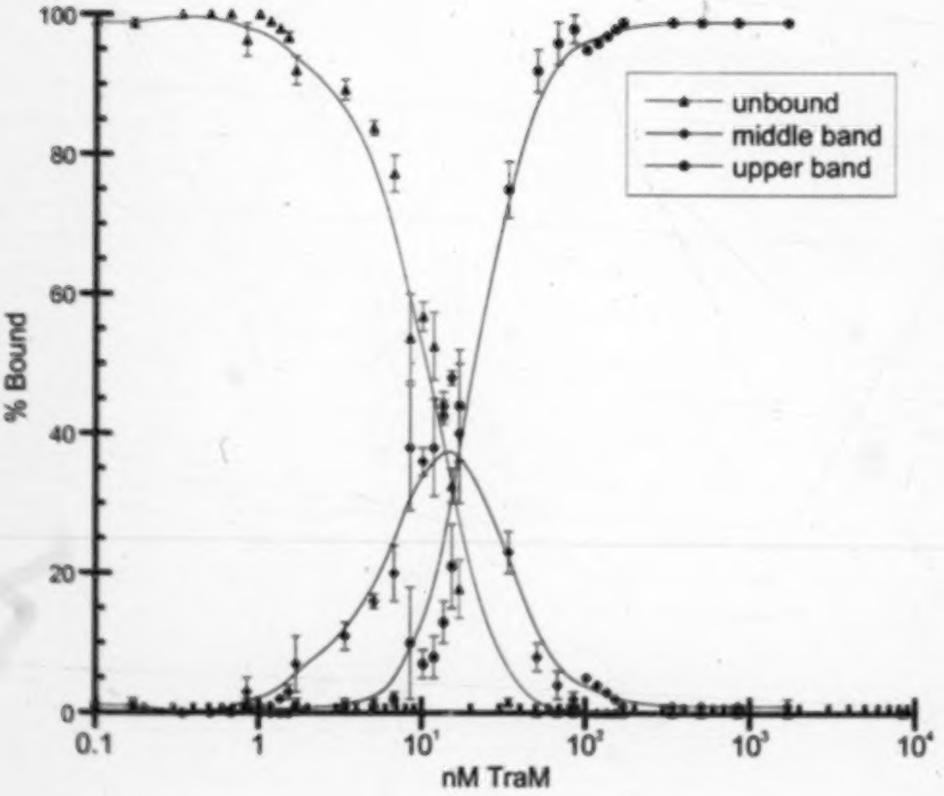

100
80
60
40
20
0
% Bound
0.1
1
10^1
10^2
10^3
10^4
nM TraM
unbound
middle band
upper band

Table 6.3

Oligonucleotides used in EMSA of TraM

Oligonucleotide	Description	Sequence
RFE12	Full *sbmA* upper strand	CTAGGATACCGCTAGGGGCGCTGCTAGCGGTGCGTC
RFE20	Full *sbmA* upper strand (less stable)	AGGATACCGCTAGGGGCGCTGCTAGCGGCGCTG
RFE50	First *sbmA* half site from upper strand	ATACCGCTAGGGGCGC
RFE54	First *sbmA* half site from upper strand with 10 bases 5' + 3'	**ATTCACGCGT**ATACCGCTAGGGGCGC**GGATCCAAGC**
RFE55	Full *sbmA* upper strand with 5 base insertion	GATACCGCTAGGGGCGCTGCTAGCGGTGCGTCC TCATA
RFE56	Full *sbmA* upper strand with 5 base deletion	CTAGGATACCGCTAGGGG-----CTAGCGGTGCGTCC
RFE11	Full *sbmA* lower strand	CTAGGGACGCACCGCTAGCAGCGCCCCTAGCGGTATC

Figure 6.8. Binding of TraM to single-strands of the *sbmA* site. EMSA of RFE11 (the upper strand of *sbmA*), RFE12 (the lower strand of *sbmA*), and RFE20 (the destabilized upper strand of *sbmA*) at various TraM concentrations (in nM) which are indicated above each lane. The intermediate band discussed in the text is shown with an arrow.

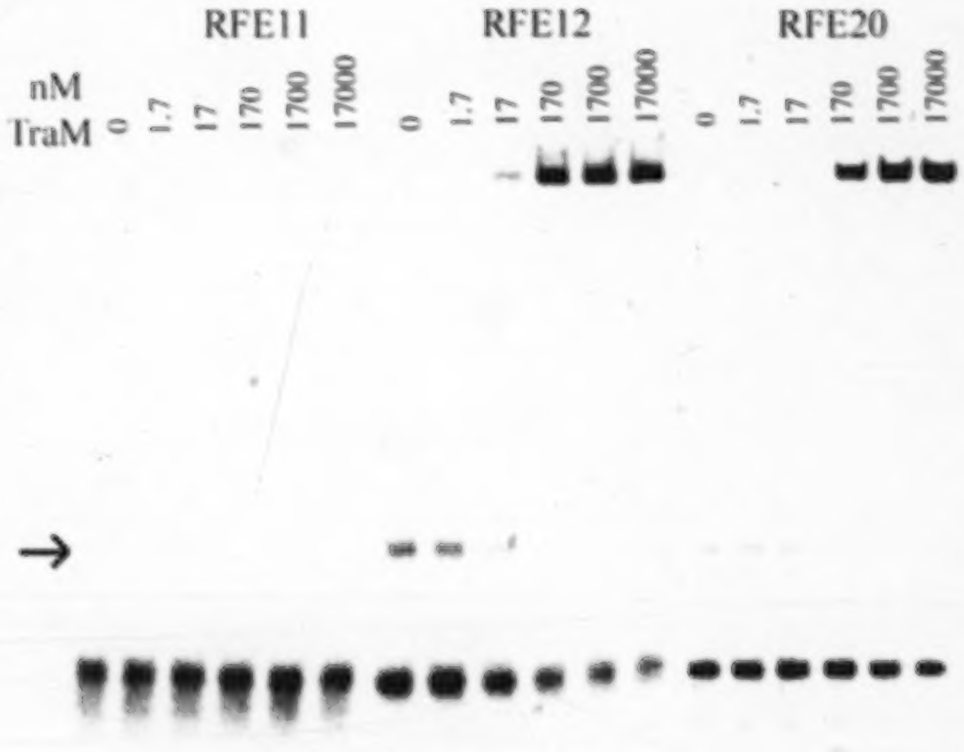
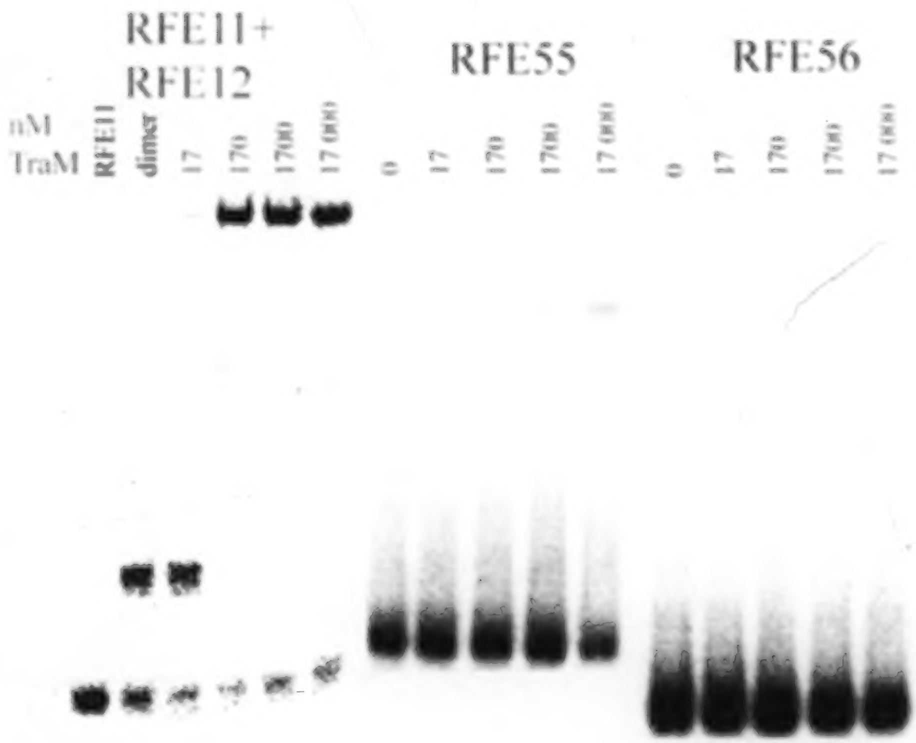

RFE11
RFE12
RFE20
nM
TraM
0
1.7
17
170
1700
17000
0
1.7
17
170
1700
17000
0
1.7
17
170
1700
17000
→

Figure 6.9. Characterization of TraM binding to a modified *sbmA* site. TraM was added in 10-fold dilutions and is shown above each lane in nM. Shown is the binding of TraM to dimerized *sbmA* single-strands and to single-stranded oligonucleotides with altered spacing between the *sbmA* half sites. RFE11 and RFE12 are the *sbmA* lower and upper strands respectively, RFE55 is *sbmA* with a 5 base insertion, and RFE56 is *sbmA* with a 5 base deletion. Dimers were made by adding RFE12 to radiolabelled RFE11.


RFE11+
RFE12
RFE55
RFE56
nM
TraM
RFE11
dimer
17
170
1700
17 000
0
17
170
1700
17 000
0
17
170
1700
17 000

look for stable secondary hairpin structures. Based on these predictions, single base changes were made which would reduce the stability of predicted hairpin structures, but maintain the homology to the consensus sequences for F TraM binding sites (discussed later). Table 6.3 shows RFE20, in which nucleotides 29, 32 and 33 were changed from T to C, G to T, and T to G respectively. These changes would decrease the proposed hairpin structure, but retain the consensus TraM binding site sequence. In Figure 6.8 it can be seen that there is significantly less of the intermediate species suggesting that this species was a hairpin structure. Affinity of TraM for this oligonucleotide was the same as for the wild-type oligonucleotide suggesting that the hairpin was not essential for binding. To analyze this further, two oligonucleotides were created which changed the spacing between the two half-sites of TraM (RFE55 and RFE56, Table 6.3). If TraM did bind a hairpin structure then increasing or decreasing the size of the loop should not be critical to binding. Figure 6.9 shows the behavior of RFE55 (RFE12 with the addition of 5 bases) and RFE56 (RFE12 with the removal of 5 bases) during EMSA. It can be seen that there is very little binding of these oligonucleotides in comparison to that of the wild-type, RFE12, or the hairpin destabilizing oligonucleotide, RFE20. This suggests that spacing between the half sites is important for TraM binding and suggests that TraM does not bind a hairpin structure.

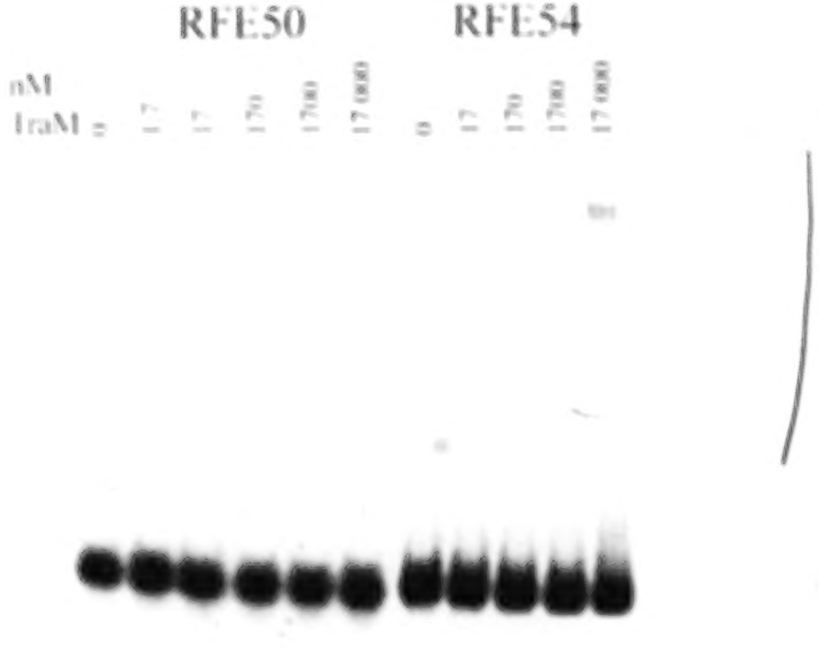
The ability of TraM to bind a single-stranded half site of *sbmA* was also of interest. If TraM did bind linear single-stranded DNA, then it should be able to bind to a single-stranded half site with a reduced affinity. This would also eliminate the possibility that secondary structure had an effect on TraM binding-single stranded DNA.

Oligonucleotides representing all four of the half sites (top and bottom strand) were synthesized and analyzed for their ability to bind TraM. RFE50, a 16 base oligonucleotide representing the first repeat in *sbmA* is shown in an EMSA in Figure 6.10. It can be seen that RFE50 is not bound by TraM, and similar results were found for all of the *sbmA* half sites. All DNA used until this point had been larger than 25 bases, while RFE50 was only 16 bases, so it is possible that binding was not seen because the protein could not properly bind to short DNA fragments. To test this, 10 bases of non-specific DNA were added to each end of each of the half site oligonucleotides. RFE54, a larger version of RFE50 (Table 6.3), bound TraM (Figure 6.10) at very high concentrations. Non-specific interactions were ruled out since RFE11 was not bound at these protein concentrations (Figure 6.8). Similar results were found for all of the half sites with extra DNA. This suggests that TraM requires DNA flanking its binding site in order to bind, and supports the previous data which suggests that TraM does bind single-stranded DNA. The binding affinities of TraM decrease in the order: double-stranded full *sbmA* site, double-stranded half site, single-stranded full site, and to the single-stranded half site.

Testing for cooperativity

There are many methods for the determination of cooperativity (Freilfelder, 1982 pp.654-684; Carlson, 1993; Chatterjee, 1996; Carey, 1991). Therefore, four different methods were chosen in order not to bias any particular aspect of cooperativity (Table 6.2):

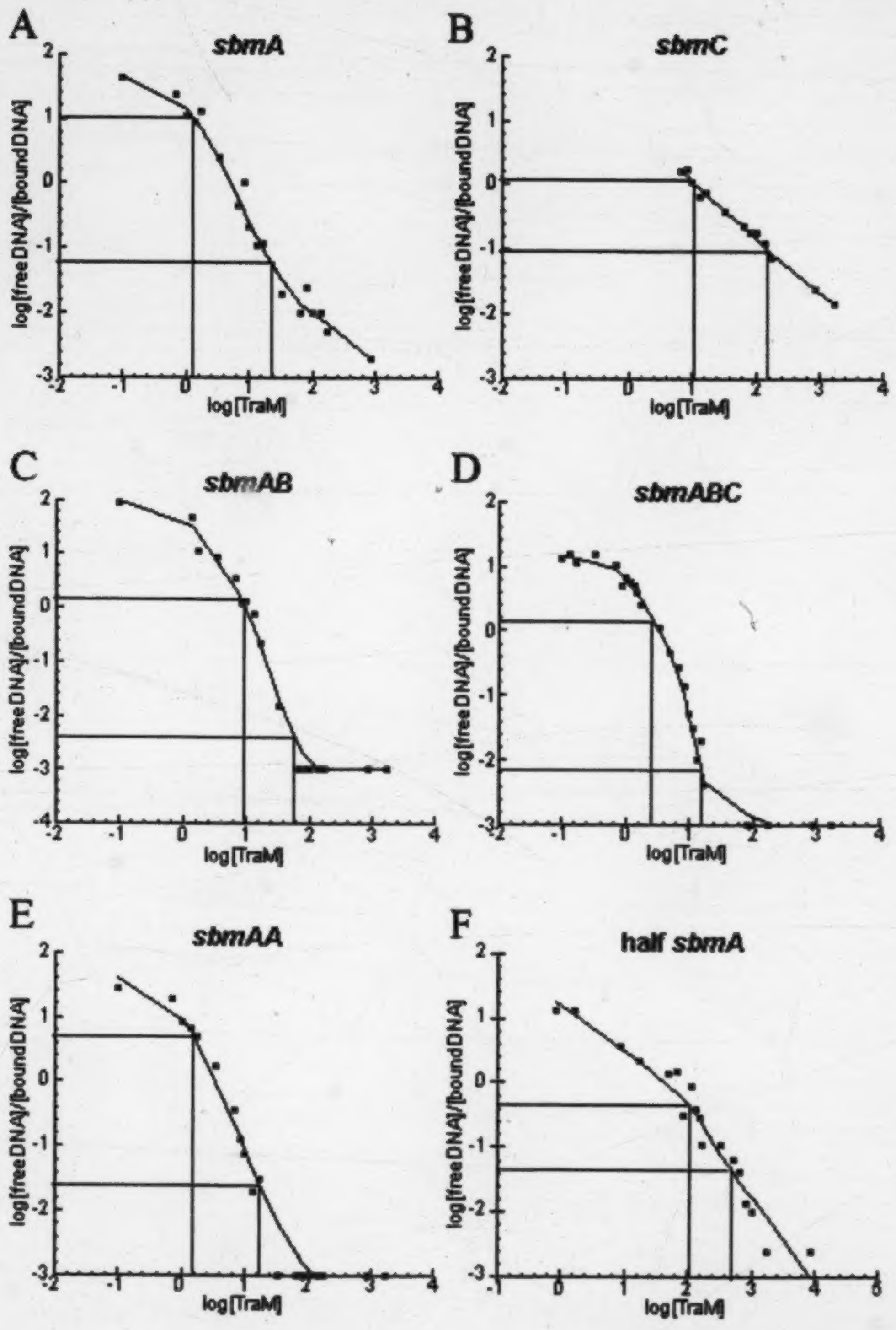
Figure 6.10. Analysis of the ability of TraM to bind *sbmA* half sites. RFE55 represents the first repeat of *sbmA* and of RFE56 represents the first repeat of *sbmA* with 10 bases added 5' and 3'. Concentrations of TraM (in nM) are indicated above each lane.


RFE50
RFE54
nM
TraM =
1.7
17
170
1700
17 000
0
17
170
1700
17 000

1) Using the basic equation for dissociation constant, Hill plots can be performed. Using this method, proteins which switch from cooperative to non-cooperative binding can be clearly distinguished. Hill plots also give a numerical degree to the cooperativity, called the Hill coefficient (n_H), which is taken from the maximum slope of the curve. When n_H is less than one, noncooperative binding is assumed. When n_H is greater than 1 then cooperative binding is assumed. The number of binding sites (n) can also be compared to n_H to give an idea of the degree of cooperativity. In a hypothetical situation where binding is completely cooperative n_H will be equal to n. This means that all sites are bound at exactly the same time, however, this situation is not found in nature. For most cooperative systems, n_H is greater than 1, but less than n. Hill plots were performed on *sbmA*, *sbmC*, *sbmAB*, *sbmABC*, *sbmAA*, and the half *sbmA* site (Figure 6.11). Hill coefficient for *sbmA* was 1.7, which is very close to n, which is 2 for *sbmA*. This indicates a highly cooperative system of binding to *sbmA*. Hill coefficient of 0.9 was found for *sbmC*, indicating non-cooperative binding. Hill coefficients of 3.25 and 2.9 were found for *sbmAB* and *sbmABC*, respectively. This indicates cooperative binding of TraM to both these fragments, and nears the presumed n of 4 for *sbmAB*. A Hill coefficient of 2.2 was found for *sbmAA*, indicating cooperative binding, and a coefficient of 1.4 was found for the half *sbmA* site. The value of the result found for the half *sbmA* site is questionable since there is only one binding site present, however, was included in the assay for interest.

2) Binding curves can be used to determine affinity with the breadth of the binding curve suggesting how strongly a molecule binds to its target. Steeper curves suggest

Figure 6.11. Hill plots of the *sbmA*, *sbmC*, *sbmAB*, *sbmABC*, *sbmAA* and half *sbmA* TraM binding sites. Calculation for Hill plots are described in Materials and Methods, and points taken to calculate the maximum slope to give n_H are shown with lines. The Y-axis represents the log of the concentrations of free DNA divided by the complexed DNA, and the X-axis represents the log of the TraM concentration.


A
sbmA
B
sbmC
C
sbmAB
D
sbmABC
E
sbmAA
F
half sbmA
log[free DNA]/[bound DNA]
log[TraM]
2
1
0
-1
-2
-3
-4
-2
-1
0
1
2
3
4
5

that there is a component of cooperativity in the binding that can result in a sigmoidal shape (Freifelder, 1982). Shallower binding curves suggest that binding occurs in a linear manner as protein concentration is increased. Using this aspect of binding curves, the breadth of the transition from 10% to 90% bound target is calculated and an increase of less than 1.81 log units suggests a cooperative mode of binding (Carey, 1991). Values greater than 1.81 suggest no cooperativity and very large values can even suggest negative cooperativity. Using this method, TraM binding to *sbmA*, *sbmAB*, *sbmABC*, and *sbmAA* is suggested to be cooperative (Table 6.2) and produce sigmoidal curves. This data also suggests that TraM binds most cooperatively to *sbmAB* and *sbmABC*. However, the degree of cooperativity is determined by fitting the data to complex theoretical curves which are precalulated for various degrees of cooperativity and for each experimental system. Small differences in affinities and the possible contribution of more complex interactions to binding can skew these theoretical curves, therefore these were not determined. Values for *sbmC* and half the *sbmA* site are greater than 1.81 and their curves are shallower than for the other sites. This suggests that there is no cooperative aspect to the binding of these sites.

3) By the definition of cooperativity, binding of one protein molecule "increases the affinity of the protein for another relatively weak binding site" (Carlson, 1993). Therefore, when 2 species are seen during EMSA, a cooperative system will have less intermediate species than a non-cooperative system and can be quantitated using band intensities (Chatterjee, 1996). To measure this the cooperativity factor, τ, was calculated for fragments which gave more than one bound species (Table 6.2). Each

of the bands for *sbmAB*, *sbmABC*, and *sbmAA* was quantitated and applied to the following equation:

$$\tau = \frac{4(\text{unbound species})(\text{fully bound species})}{(\text{intermediate species})^2}$$

Cooperative interactions can be assumed if values of τ exceed 1. Using this assay all three fragments are suggested to display cooperative interactions, with *sbmABC* having the greatest value.

4) The last method for measuring cooperativity is based on the value of ω (Carlson, 1993) which can be determined if the association constants of the sites within a larger site are known. For example if a site is composed of 2 sites, X and Z:

$$\omega = \frac{(1/K_a \text{ of } X)(1/K_a \text{ of } Z)}{(\text{free protein at 50\% binding of } X \text{ in } XZ)(\text{free protein at 50\% binding of } Z \text{ in } XZ)}$$

In these analyses both of the smaller sites were identical, therefore:

$$\omega = \frac{(1/K_a \text{ of smaller site})^2}{(\text{free protein at 50\% binding of larger site})^2}$$

The amount of DNA in the reactions is insignificant, so the amount of bound protein is also insignificant. Therefore the total amount of protein at 50% binding was used in the lower part of the equation. In this analysis values of ω greater than 1 indicate cooperativity. Calculations for ω were made for *sbmA*, *sbmAB*, and *sbmAA* and are shown in Table 6.2. Calculations were made for *sbmA* since the K_a of the half site was known, for *sbmAA* since the K_a for *sbmA* was known, and for *sbmAB* since this is basically two *sbmA* sites in a direct repeat orientation. Calculations were not made for *sbmABC*, even though the K_a are known for all the sites, since the presence of 3 sites complicates the equation. Using this analysis TraM appears to bind to *sbmA* with a great deal of cooperativity. Binding to *sbmAB* and *sbmAA* is also suggested to be cooperative, although to a lesser extent.

DNA bending

The fact that molecules with a large cross sectional radius migrate more slowly than smaller molecules was used to determine if TraM bends DNA upon binding. A DNA molecule with a bend at its end will have almost the same end-to-end distance as the linear molecule (L), however a DNA molecule of the same length bent in the center will have an end-to-end distance represented by $L[\cos(\alpha/2)]$ where α is the bend angle of the DNA (Kim *et al.*, 1989). Therefore $\mu_M/\mu_E = L[\cos(\alpha/2)] / L = \cos(\alpha/2)$ where μ_M is the mobility of the complex with the binding site in the center, and μ_E is the mobility of the complex with the binding site at the end. However, it should be noted that since factors other than end-to-end distance effect mobility (see below for examples), the calculated bend angles may be different from absolute bend angles. Bend angles for *sbmA*, *sbmC*, *sbmAB* and *sbmABC* were calculated using this equation. PCR products using primers RFE2 and RFE3 (annealing inside the multiple cloning site of pBEND2, Figure 6.1, and amplifying the cloned binding sites) were cut with a variety of restriction enzymes and run with and without TraM on 8% polyacrylamide gels. Concentrations of TraM were used which gave fully bound species. TraM at a concentration of 8.5 nM was used for *sbmA* (Figure 6.12), 170 nM for *sbmC* (Figure 6.12), 10 nM and 70 nM for *sbmAB* (Figure 6.13), and 17 nM for *sbmABC* (Figure 6.13). DNA cut with *Bam*HI was used as the template with the binding site at the end, and DNA cut with *Eco*RV (or *Dra*I for *sbmA*) was used as the template with the binding site in the middle. Using this analysis,

Figure 6.12. Analysis of TraM's ability to bend *sbmA* and *sbmC* upon binding. A. EMSA of *sbmA* cut with a variety of restriction enzymes (designated above each lane). 8.5 nM TraM was used to bind *sbmA* and the unbound and bound binding sites are designated by the lower and upper arrows, respectively. B. EMSA of *sbmC* cut with a variety of restriction enzymes (designated above each lane). 170 nM TraM was used to bind *sbmC* and the unbound and bound binding sites are designated by the lower and upper arrows, respectively. C and D. Graphs representing the bending of *sbmA* (C) and *sbmC* (D) bound by TraM. Relative mobilities (Rf; Y-axis) were calculated by comparing the mobility of the retarded fragments to the unbound DNA species. X-axis represents the number of base pairs from one side of each fragment to the center of the binding site (one side was arbitrarily chosen).

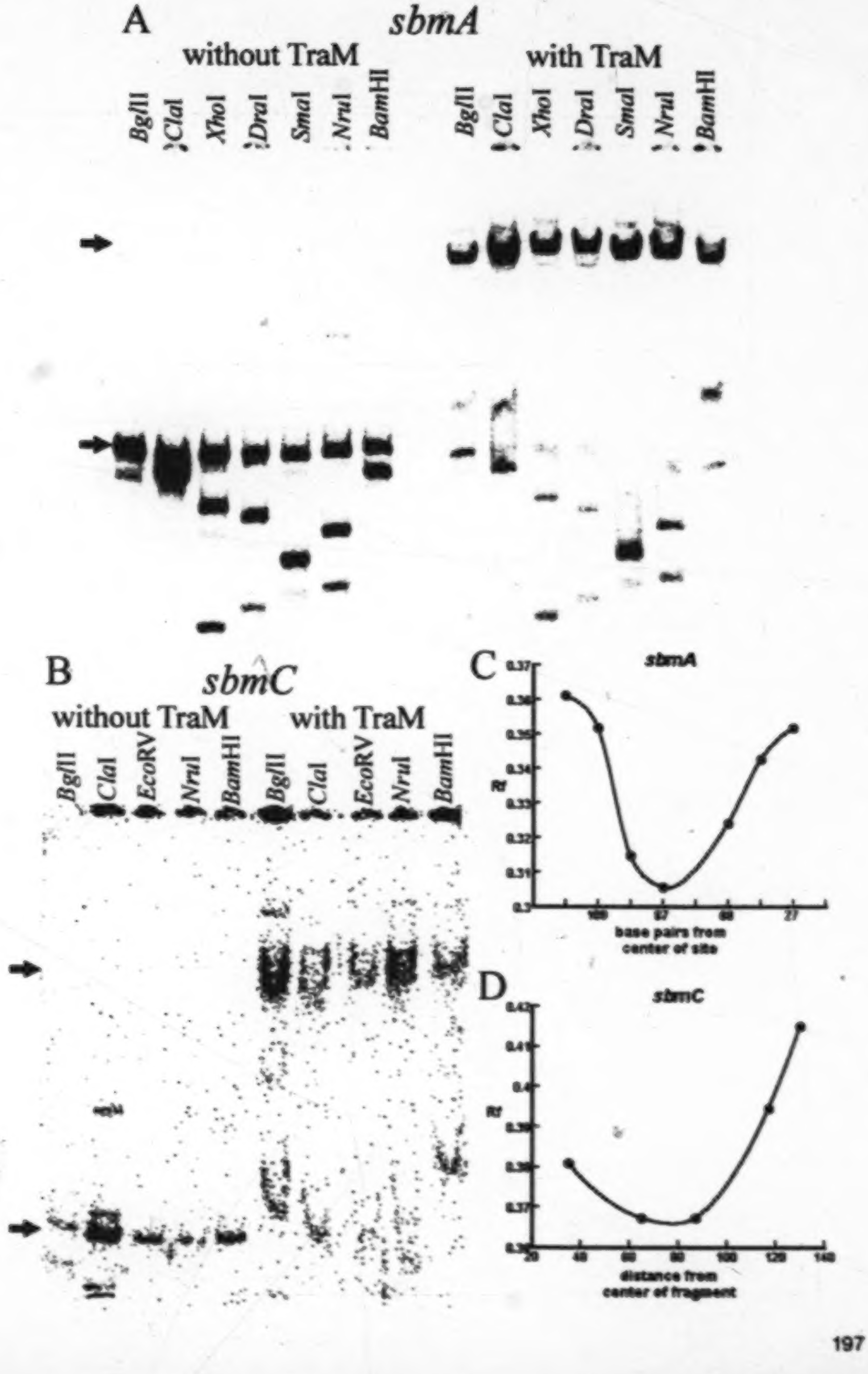
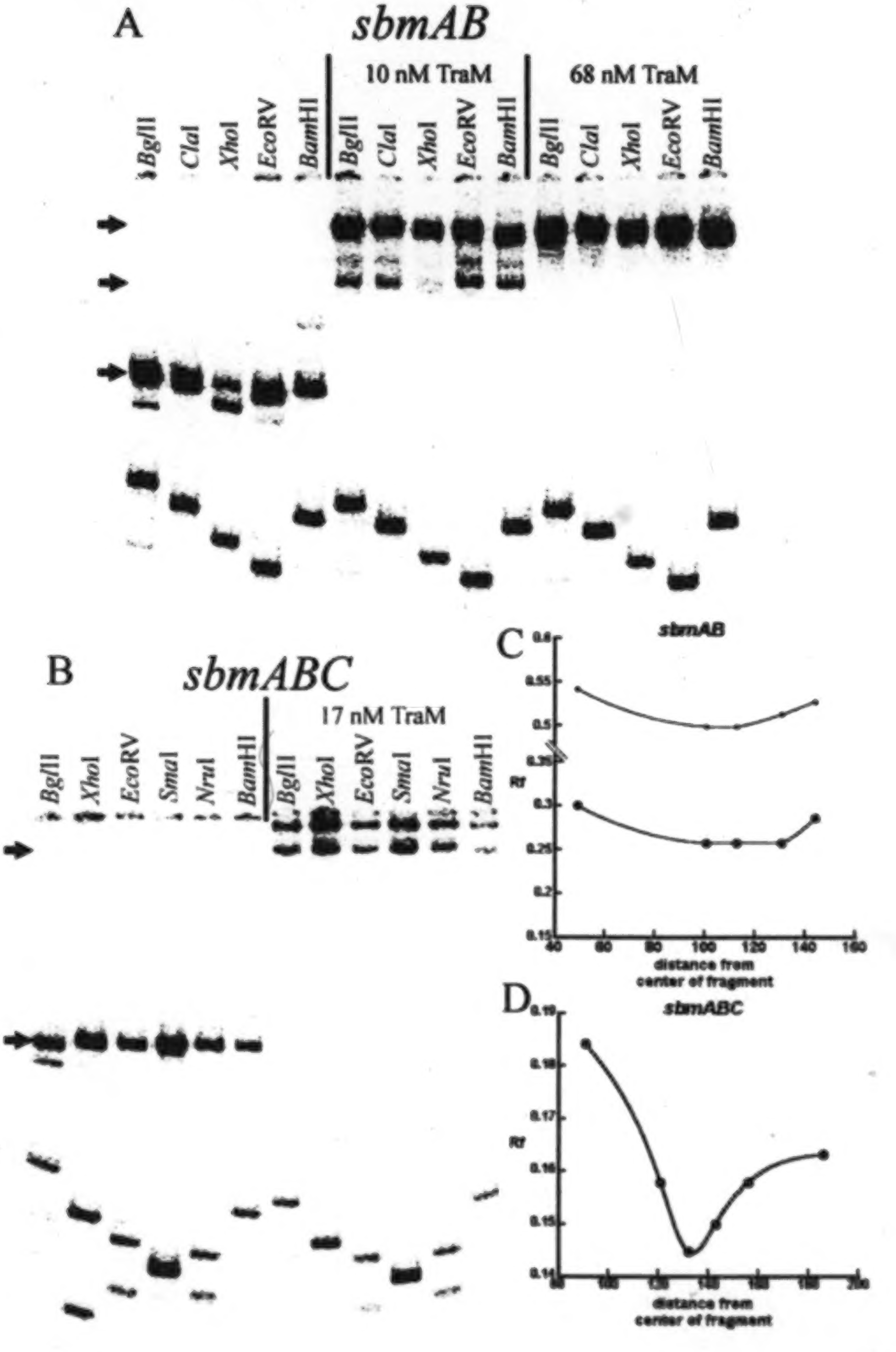

A
sbmA
without TraM
with TraM
BglII
ClaI
XhoI
DraI
SmaI
NruI
BamHI
BglII
ClaI
XhoI
DraI
SmaI
NruI
BamHI
B
sbmC
without TraM
with TraM
BglII
ClaI
EcoRV
NruI
BamHI
BglII
ClaI
EcoRV
NruI
BamHI
C
sbmA
0.37
0.36
0.35
0.34
0.33
0.32
0.31
0.3
Rf
109
87
60
27
base pairs from
center of site
D
sbmC
0.42
0.41
0.4
0.39
0.38
0.37
0.36
Rf
20
40
60
80
100
120
140
distance from
center of fragment

Figure 6.13. Analysis of TraM's ability to bend *sbmAB* and *sbmABC* upon binding. A. EMSA of *sbmAB* cut with a variety of restriction enzymes (designated above each lane). 10 and 70 nM TraM was used for *sbmAB* and the unbound, intermediate complex, and upper complex are designated with arrows. B. EMSA of *sbmABC* cut with a variety of restriction enzymes (designated above each lane). 17 nM TraM was used for *sbmABC* and the unbound and fully bound complexes are designated with arrows. C and D. Graphs representing the bending of *sbmAB* (C; upper and lower complexes represented by red and blue complexes, respectively) and *sbmABC* (D) bound by TraM. Relative mobilities (Rf; Y-axis) are calculated by comparing the distance migrated by the retarded fragments to the unbound DNA species. X-axis represents the number of base pairs from one side of each fragment to the center of the binding site (one side was arbitrarily chosen).


A
sbmAB
10 nM TraM
68 nM TraM
BglII
ClaI
XhoI
EcoRV
BamHI
BglII
ClaI
XhoI
EcoRV
BamHI
BglII
ClaI
XhoI
EcoRV
BamHI
B
sbmABC
17 nM TraM
BglII
XhoI
EcoRV
SmaI
NruI
BamHI
BglII
XhoI
EcoRV
SmaI
NruI
BamHI
C
sbmAB
Rf
0.6
0.55
0.5
0.35
0.3
0.25
0.2
0.15
40
60
80
100
120
140
160
distance from center of fragment
D
sbmABC
Rf
0.19
0.18
0.17
0.16
0.15
0.14
80
100
120
140
160
180
200
distance from center of fragment

sbmA is bent to approximately 50°, *sbmC* to approximately 25°, *sbmAB* to approximately 25° in its primary complex and then 45° in its secondary and final complex, and *sbmABC* is bent to approximately 30° in its final complex. This suggests that bound TraM does bend *sbmA* but has a lesser effect on *sbmC*. This is presumably due to the differences in the arrangement of the binding sites and may be due to the interaction of the subunits of the bound protein. Another interesting point is that if *sbmB* is also bent by TraM, then TraM bound to *sbmA* and *sbmB* would bend DNA approximately 100°. This may facilitate complex formation where a tetramer of TraM binds to both *sbmA* and *sbmB*. The change in the calculated bend angles of bound *sbmAB* using high and low concentrations of TraM suggests that there is a difference in the conformation of the complexes. The data also shows that TraM bound to *sbmABC* does not bend DNA to the same extent as *sbmAB*. This may be due to the addition of *sbmC* and associated TraM may straighten the DNA on the *sbmC* side of the *sbmAB* binding sites, or as discussed below, the increase in the size of the cloned binding site may have a negative effect on the assay.

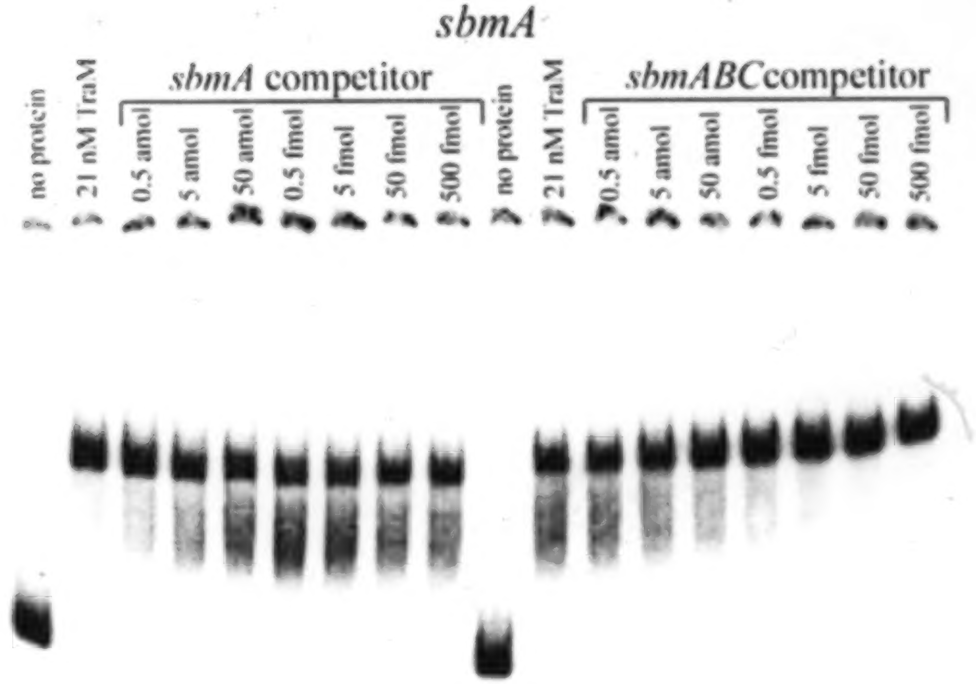
Factors other than end-to-end distance can have an effect on the mobility of complexes during electrophoresis. One factor to consider is the size of the cloned sequences. Analyzing binding sites that are 150 base pairs or larger may not be possible since excess cloned DNA may mask the effect of varying the size of the DNA on either side of the bound protein. Another factor to consider is that flanking DNA may be more important for larger complexes than for smaller complexes. In these cases varying the size of the DNA on either side of the complex may result in different protein-DNA complexes being formed. Finally all proteins analyzed using this procedure bend their targets to a maximum of 112° (Kim, 1989). If DNA is wrapped around a protein complex in *sbmAB* or *sbmABC*, the effect of varying the size of flanking DNA may not have the same effect as it does for proteins which bend their targets to only 90°. For this reason the data for multiple binding sites is not as clear as it is for single binding sites such as *sbmA* and *sbmC*.

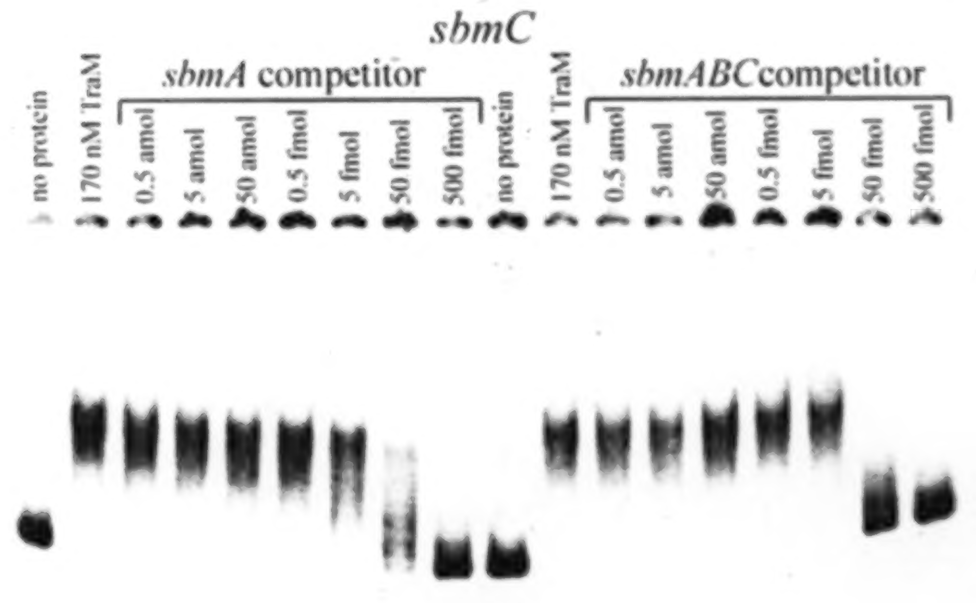
Competition assays

In order to determine the stability of TraM bound to its binding sites, competition assays were performed using 0.5 fmol of radiolabelled template and incubated with various amounts of TraM for 15 minutes at 37°C. 0.5 amol to 500 fmol of non-radioactive *sbmA* and *sbmABC* were then added and incubated for ~~an~~ additional 15 minutes at 37°C. Figure 6.14a shows competition reactions of *sbmA* and *sbmC* using 21nM and 170 nM TraM respectively. Neither *sbmA* nor *sbmABC* were able to compete TraM from *sbmA* very well. At 1000 times excess competitor, some competition was seen, however, this weak level of competition did not allow for the calculation of dissociation constants. The fact that the dissociation and association constants are so different suggests that TraM stably binds to its binding sites. Smearing was seen below the bound species, but this was an artifact of the gel. The fact that more smearing was seen in the center lanes suggests that this may have resulted from the buildup of heat while running the gel.

A competition assay of *sbmC* is shown in Figure 6.14b using 170 nM of TraM. 10-fold excess *sbmA* and *sbmABC* are required to begin to compete TraM from *sbmC*. At 100-

Figure 6.14. EMSA competition assays to test the stability of TraM bound to *sbmA* and *sbmC*. 21 nM TraM was used for *sbmA* and 170 nM for *sbmC*. 0.5 fmol of radioactive template was used in each lane with the amount and type of competitor specified above each lane. The position of unbound fragments are shown in the lanes with no protein.


sbmA
no protein
21 nM TraM
sbmA competitor
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol
no protein
21 nM TraM
sbmABC competitor
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol

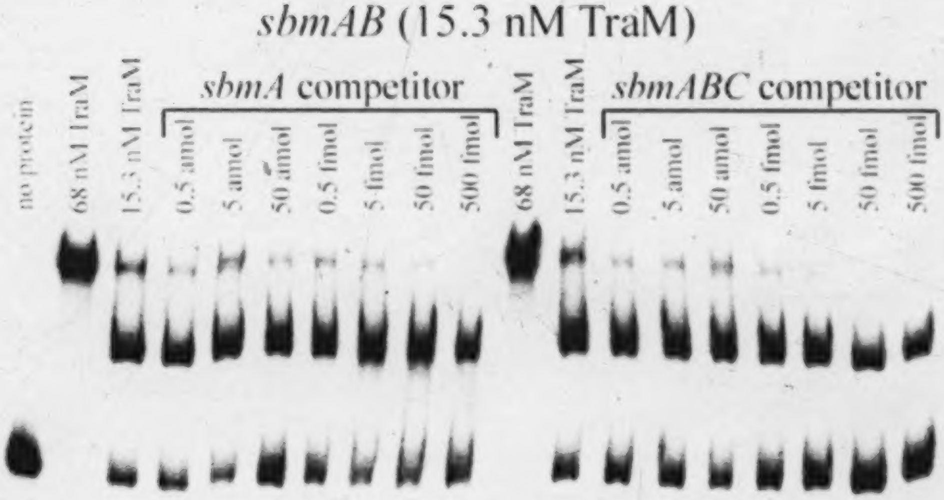

sbmC
no protein
170 nM TraM
sbmA competitor
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol
no protein
170 nM TraM
sbmABC competitor
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol

fold excess of competitor, *sbmABC* completely competes TraM from *sbmC*, whereas *sbmA* does not, suggesting that *sbmABC* is a better competitor than *sbmA* (presumably since *sbmABC* has a higher affinity for TraM).

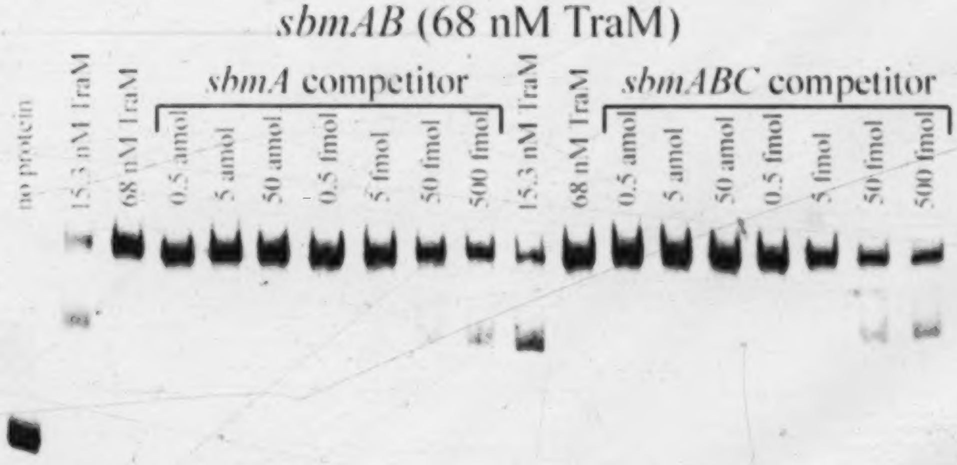
Competition assays for *sbmAB* and *sbmABC* were done using 2 concentrations of TraM (Figure 6.15 and 6.16 respectively). This was done in order to look at the stability of each of the 2 species of bound DNA shown in the retardations. Therefore, 15.3 nM and 68 nM TraM was used to bind *sbmAB* and competition was performed using 0.5 amol to 500 fmol of *sbmA* or *sbmABC* (Figure 6.15a and b). At 15.3 nM TraM competition of the upper band is seen (Figure 6.15a) with minor changes in the intensity of the lower retarded band. In Figure 6.15b again the upper band is competed by both *sbmA* and *sbmABC*, however unbound DNA does not accumulate. This suggests that the first bound complex is more stable than the upper complex. In both 6.15a and b, *sbmABC* is a better competitor than *sbmA*, presumably due to the higher association constant of *sbmABC*.

Competition of TraM bound to *sbmABC* was done using 6.8 nM and 17 nM TraM (Figure 6.16a and b respectively). In Figure 6.16a with 6.8 nM of TraM (competition of the lower retarded complex) there is no significant competition by *sbmA* or *sbmABC* of the lower retarded complex. Some smearing is seen with higher levels of *sbmA* competitor, but this is not significant since the intensity of the lower retarded complex does not change. Retardation in the presence of 17 nM TraM gives only upper complex and no lower complexes (Figure 6.16b). Competition does occur for TraM bound to *sbmABC* in

Figure 6.15. EMSA competition assays to test the stability of TraM bound to *sbmAB*. TraM was used at 15.3 nM and 68 nM and examples of both protein concentrations are shown in both gels. 0.5 fmol of radioactive template was used in each lane with the amount and type of competitor specified above each lane. The position of unbound fragments are shown in the lanes with no protein.


sbmAB (15.3 nM TraM)

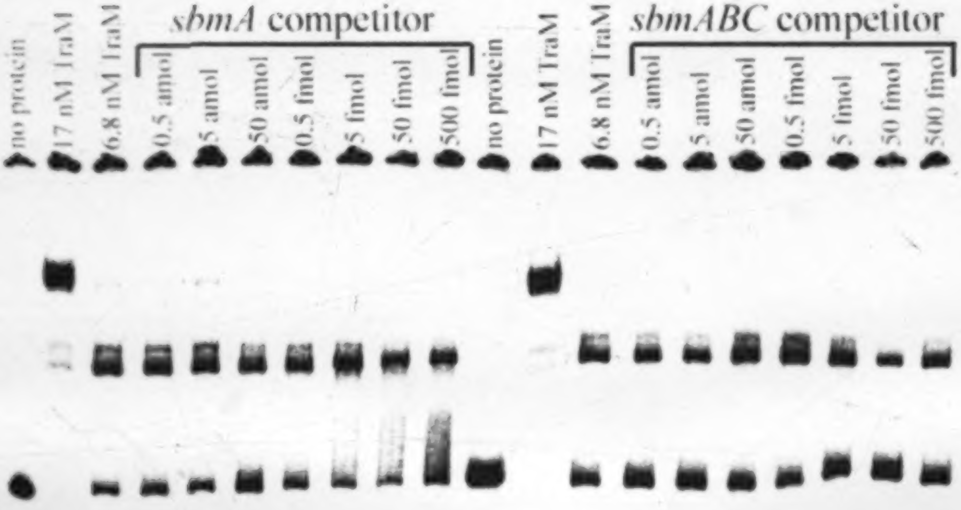
no protein
68 nM TraM
15.3 nM TraM
sbmA competitor
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol
68 nM TraM
15.3 nM TraM
sbmABC competitor
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol


sbmAB (68 nM TraM)

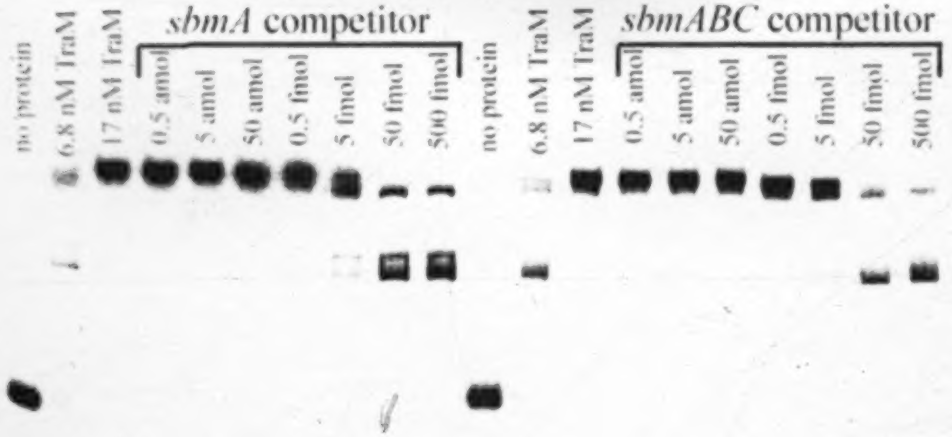
no protein
15.3 nM TraM
68 nM TraM
sbmA competitor
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol
15.3 nM TraM
68 nM TraM
sbmABC competitor
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol

Figure 6.16. EMSA competition assays to test the stability of TraM already bound to *sbmABC*. TraM was used at 6.8 nM and 17 nM and examples of both protein concentrations are shown in both gels. 0.5 fmol of radioactive template was used in each lane with the amount and type of competitor specified above each lane. The position of unbound fragments are shown in the lanes with no protein.

sbmABC (6.8 nM TraM)


sbmA competitor
sbmABC competitor
no protein
17 nM TraM
6.8 nM TraM
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol
no protein
17 nM TraM
6.8 nM TraM
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol

sbmABC (17 nM TraM)

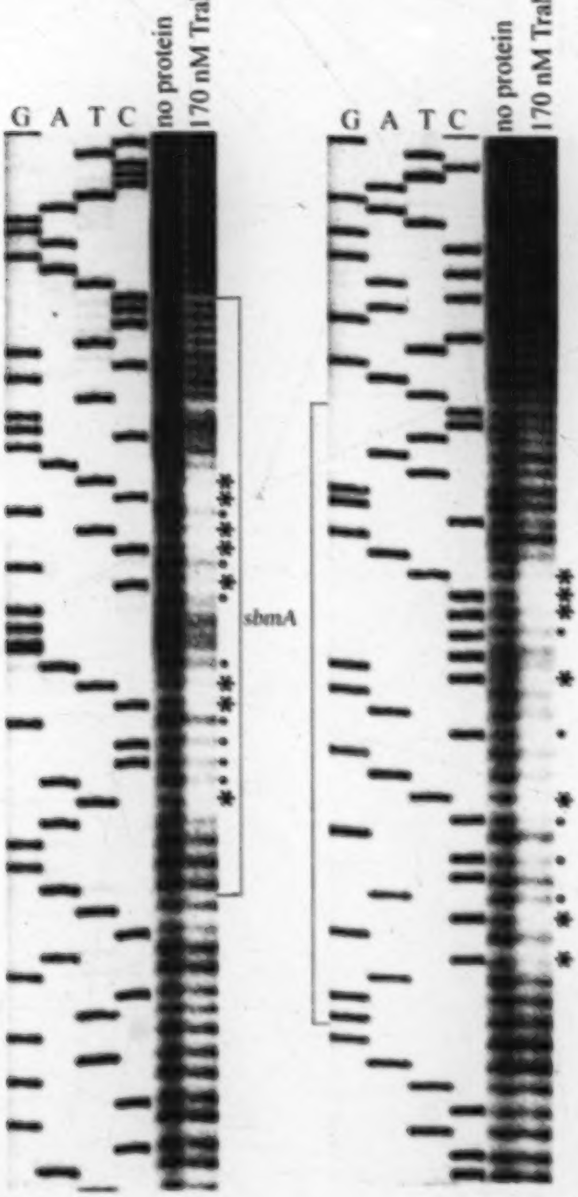

sbmA competitor
sbmABC competitor
no protein
6.8 nM TraM
17 nM TraM
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol
no protein
6.8 nM TraM
17 nM TraM
0.5 amol
5 amol
50 amol
0.5 fmol
5 fmol
50 fmol
500 fmol

the upper complex in the presence of 5, 50, and 500 fmol of *sbmA* and *sbmABC*. This is similar to *sbmAB* and suggests that the upper complex is not as stable a species as the lower bound complex.

Hydroxyl radical footprinting

To determine which nucleotides are protected in the TraM binding sites, hydroxyl radical footprinting of *sbmA* and *sbmC* was performed. Hydroxyl radicals are able to cleave the DNA backbone by abstraction of a hydrogen atom bound to a carbon atom. The cleavage reaction occurs preferentially at the 5' and 4' hydrogen atoms (Balasubramanian *et al.*, 1998). The cleavage is not base-specific, which allows footprinting to occur irrelevant of the DNA sequence. Cleavage was performed on DNA bound by TraM which had been amplified using PCR with one radioactively labeled primer. Footprinting of the upper and lower strands of *sbmA* is shown in Figure 6.17. The first 4 lanes are a sequencing reaction of the binding site using the same primer that was radioactively labeled in the footprinting reaction. Strongly protected bases are shown with a large asterisk and those that are weakly protected are shown with a small asterisk. Using WebLabViewer Light® the protected bases were mapped onto a model of a double-stranded DNA fragment in the β-form. Figure 6.18A and B shows two sides of the DNA molecule with red and pink bases representing strongly and weakly protected bases on the upper strand respectively, and blue and light blue bases representing strongly and weakly protected bases on the lower strand. The identity of the strongly protected bases are also shown roughly above or below their position in the molecule. The radioactively labeled bases are shown as a gold ring with a red or blue cap representing the top and bottom strands. Bases on the

Figure 6.17. Hydroxyl radical footprinting of the top and bottom strands of *sbmA*. Sequencing reactions using the same primers as those used in the footprinting are shown to the left of each of the footprinting lanes. The *sbmA* binding site is shown in the center of the figure with boxes defining the size of the DNase I footprint. Large and small asterisks show the strongly and weakly protected bases respectively. Protein concentrations used in the experiment are shown above the lanes.



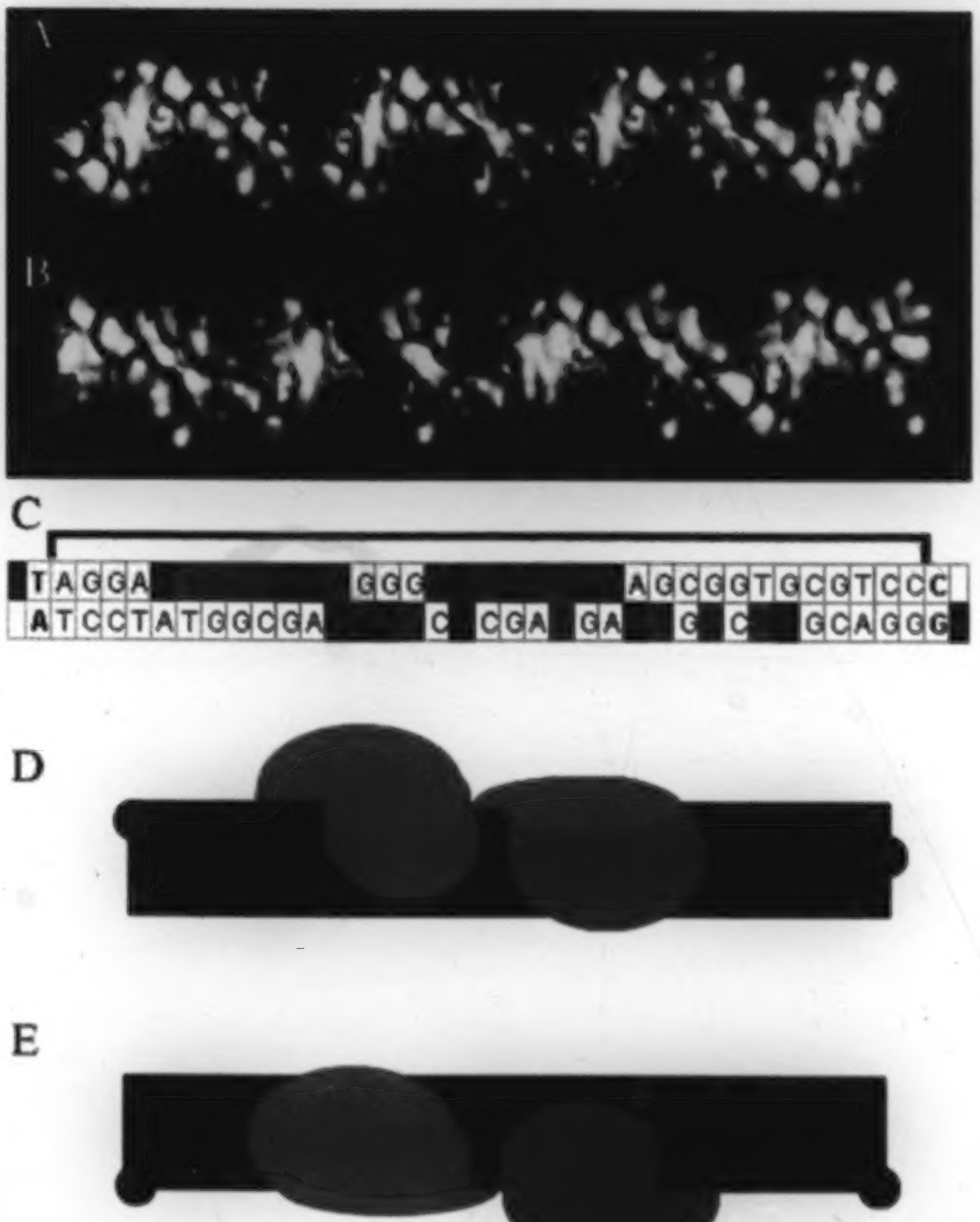
sbmA

Top strand

Bottom strand

G A T C
no protein
170 nM TraM
sbmA
G A T C
no protein
170 nM TraM

Figure 6.18. Mapping the protected bases on *sbmA* and proposal of a model for TraM binding. A and B. Bases protected by TraM during hydroxyl radical footprinting are mapped onto a double-stranded DNA helix. The strongly and weakly protected bases on the upper strand are shown in red and pink respectively. The strongly and weakly protected bases on the lower strand are shown in dark and light blue respectively. The 5' phosphate labels are shown at the ends of the helix colored red and blue for the upper and lower strands, respectively, with a ring of gold below them. Identities of the strongly protected bases are also designated in red and blue for the upper and lower strands, respectively. A shows the view of one side of *sbmA* and B shows another view rotated approximately 90°. C. The sequence of the *sbmA* site is shown with the strongly and weakly protected bases color coded as in A and B. The DNase I footprint of the *sbmA* site is also shown above the sequence. D and E. A model of the DNA binding portion of two TraM monomers on the DNA. Views of D and E are shown to represent the views seen in A and B respectively. Positions of the 5' labels are shown in red and blue for orientation.


A
B
C
TAGGA GGG AGCGGTGCGTCCC
ATCCTATGGCGA C CGA GA G C GCAGGG
D
E

sides of these pictures were not protected and are shown in white. In Figure 6.17C the sequence of the *sbmA* site is shown with strongly protected bases darkly shaded and weakly protected bases lightly shaded. It is interesting that bases before the conserved CATGGGG motif are protected. When aligning the half TraM binding sites (discussed later) this motif is conserved and the preceding sequences, which are highly protected, are not conserved. Using these methods of representation it can be hypothesized that two TraM molecules bind to the two center major groves in 6.18A and B and protect the upper and lower strands by partially wrapping around the DNA. Models for a dimer binding to the *sbmA* site is shown diagrammatically in Figure 6.18D and E and correspond to the faces shown in 6.18A and B respectively.

Hydroxyl radical footprinting of the upper and lower strands of *sbmC* was also performed and is shown in Figure 6.19. Strongly and weakly protected bases are designated with large and small asterisks respectively. The last four lanes of each figure are a sequencing reaction of the cloned *sbmC* site using the same primer that was end-labeled during the amplification of DNA for footprinting. Just as for *sbmA* in Figure 6.18A and B, the protected bases were mapped onto a DNA fragment in the β-form and is shown in Figure 6.20A and B. Since the *sbmC* site is larger it could not be mapped using a single molecule, therefore, the two footprinted regions (on opposite ends and opposite sides of the *sbmC*) are shown in two parts (6.20A and B). Again the radioactively labeled nucleotide is represented with a gold ring with a red or blue cap, representing the top and bottom strand respectively. The two molecules can be put together since the T (shown in blue) at the right of Figure 6.20A is the same base as the last T in the GTGT run (shown

Figure 6.19. Hydroxyl radical footprinting of the top and bottom strands of *sbmC* using 1700 nM TraM. Sequencing reactions using the same primers as those used in the footprinting are shown to the right of each of the footprinting lanes. The *sbmC* binding site is shown in the center of the figure with boxes defining the size of the DNase I footprint. Large and small asterisks show the strongly and weakly protected bases respectively. Protein concentrations used in the experiment are shown above the lanes.

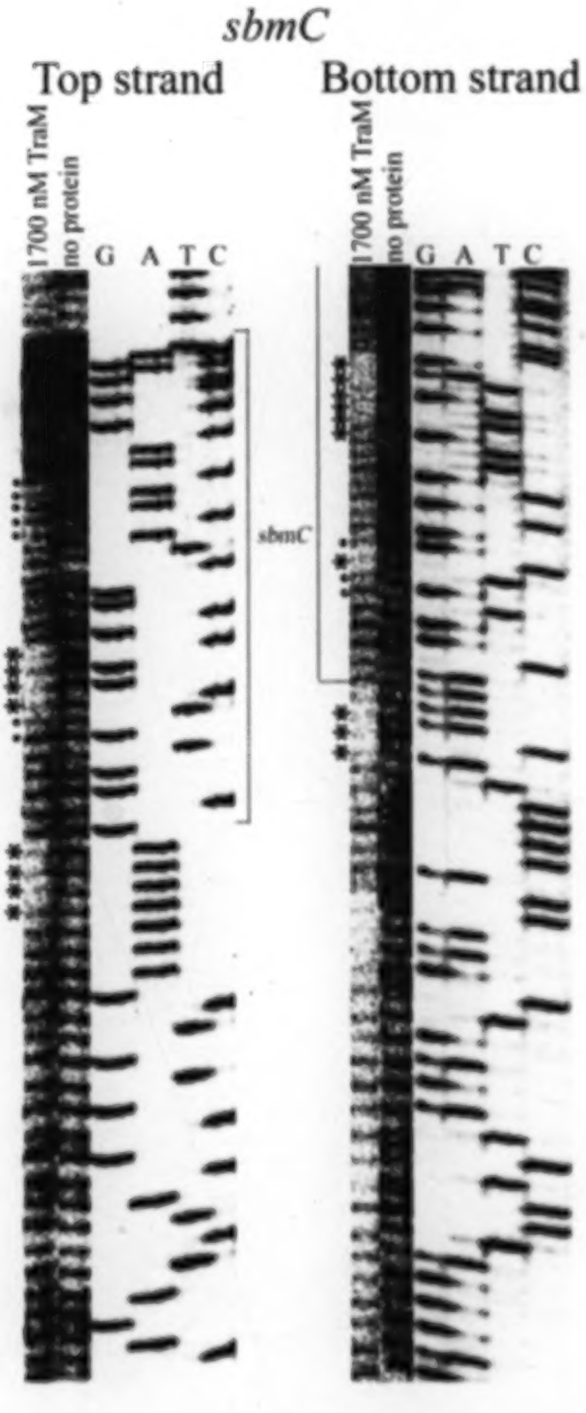
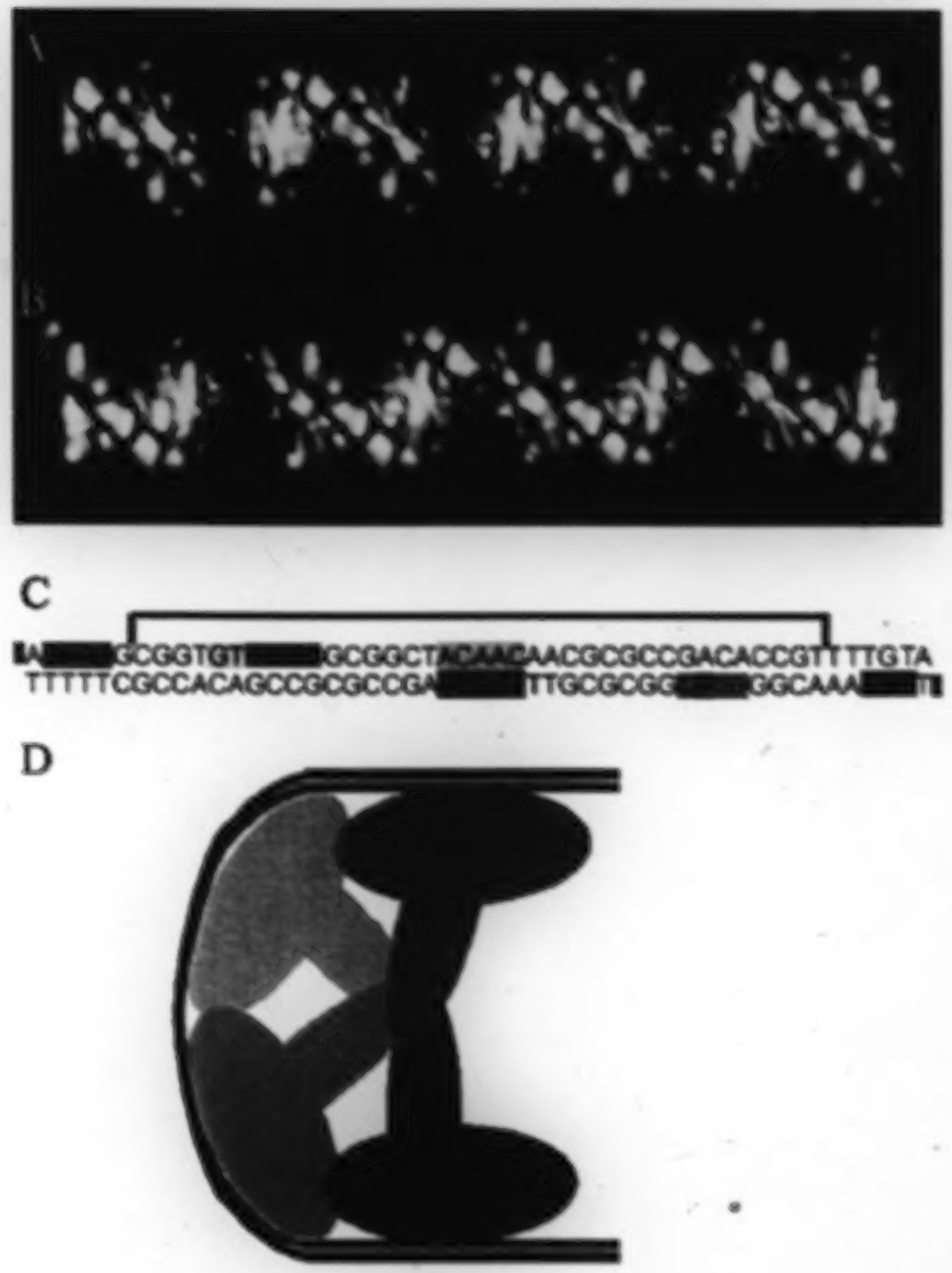

sbmC
Top strand
Bottom strand
1700 nM TraM
no protein
G A T C
1700 nM TraM
no protein
G A T C
sbmC

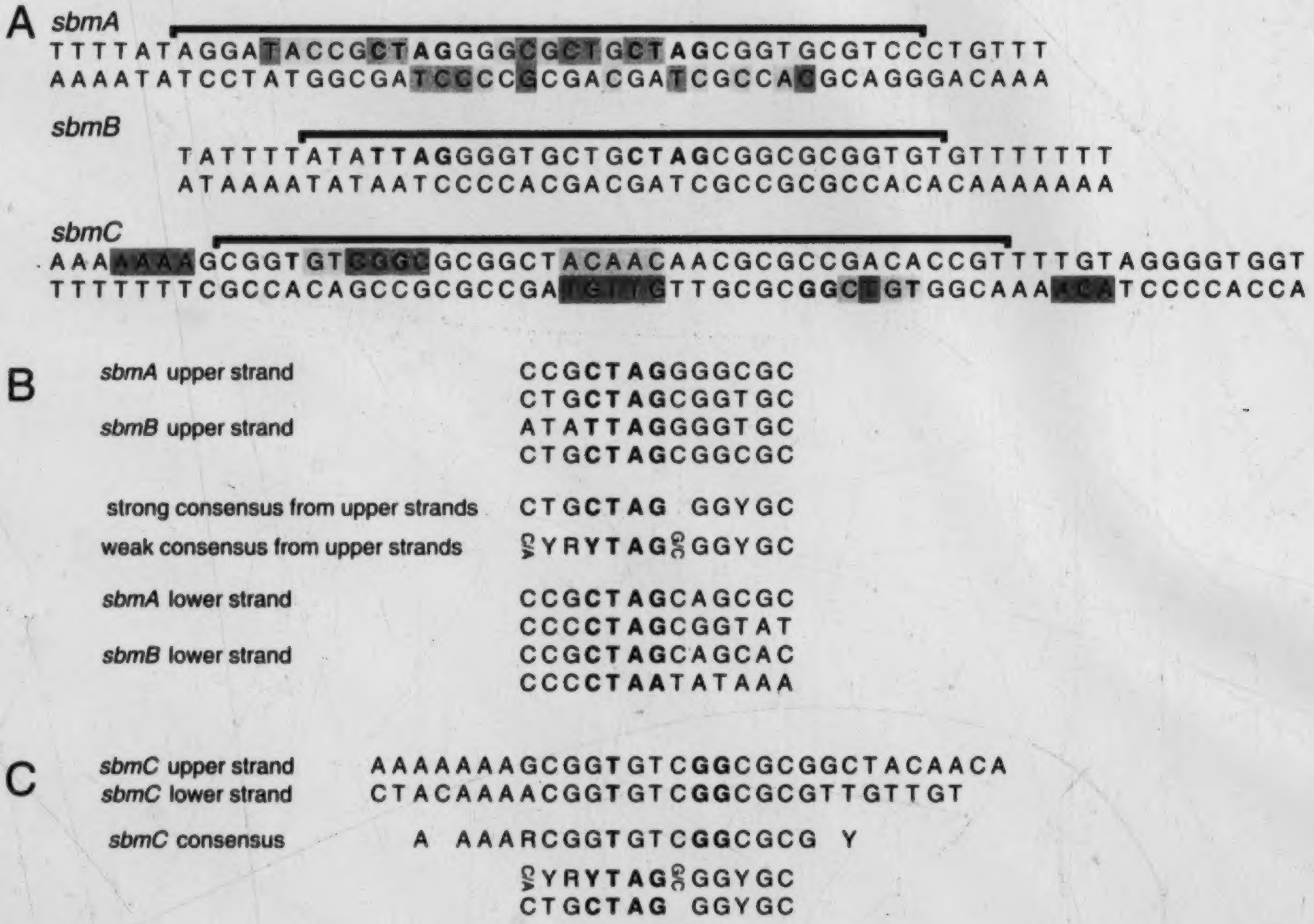
Figure 6.20. Mapping the protected bases on *sbmC* and proposal of a model for TraM binding. A and B. Bases protected by TraM during hydroxyl radical footprinting are mapped onto a double-stranded DNA helix. The strongly and weakly protected bases on the upper strand are shown in red and pink respectively. The strongly and weakly protected bases on the lower strand are shown in dark and light blue respectively. The 5' phosphate labels are shown at the ends of the helix colored red and blue for the upper and lower strands, respectively, with a ring of gold below them. Identities of the strongly protected bases are also designated in red and blue for the upper and lower strands, respectively. A shows one end of the *sbmC* site and B is a continuation of the model in A, however, is rotated 180° to the view shown in A. The blue T and G at the right side of A are the first T and the last G in the TGTTG string seen at the left of B. C. The sequence of the *sbmC* site is shown with the strongly and weakly protected bases color coded as in A and B. The DNase I footprint of the *sbmC* site is also shown above the sequence. D. A model for the binding of a TraM tetramer to the *sbmC* site.


A
B
C
A[illegible]GCGGTGT[illegible]GCGGCTACAACAACGCGCCGACACCGTTTTGTA
TTTTTCGCCACAGCCGCGCCGA[illegible]TTGCGCGG[illegible]GGCAAA[illegible]T
D

in blue) in Figure 6.20B, and the G (shown in blue) at the right of 6.20A is the same as the first G in the GTGT run (shown in blue) in Figure 6.20B. The sequence of the *sbmC* site is shown in Figure 6.20C with the strongly and weakly footprinted bases strongly and lightly shaded respectively. It is interesting that the footprint does appear to have symmetry along a central axis at the ACAAC sequence of the upper strand. Creating a model for TraM binding is not as easy as for *sbmA* since the footprint on each end of the DNA molecule spans 3 major grooves. TraM monomers may be bound to the terminal two major grooves in *sbmC* on each end. Tetramerization of the bound proteins may bend and wrap the DNA around the protein complex so that the bases in the center of the site are protected non-specifically.

Alignments of the TraM binding sites were done in order to try and determine a consensus sequence (Figure 6.21). In the first part of the figure (A) the three binding sites are shown with the DNase I footprint (solid line over the sequences), and hydroxyl radical footprints (as shown in 6.18B and 6.20B with the strongly and weakly protected bases darkly and lightly shaded respectfully). Each site also has bases which are bolded, however, these are only to provide orientation when looking at the consensus sequences. The *sbmA* and *sbmB* sites are aligned to show that the consensus sequences are equally spaced in both sites. In 6.21B, the upper strand half sites are aligned and a consensus sequence is generated. R represents purines, and Y represents pyrimidines. A strong consensus is generated which represents sequences present in 3 or more of the sites. A weak consensus is also generated representing all of the bases in the four sites. Below the consensus sequences are the half sites from the lower strands of *sbmA* and *sbmB*.

Figure 6.21. A. Sequences of the *sbmA*, *sbmB*, and *sbmC* sites in the F plasmid. Hydroxyl radical footprints of *sbmA* and *sbmC* from Figures 6.17 and 6.19 are shown in dark and light gray to represent strongly and weakly protected bases. DNase I footprints of the three sites are shown above each sequence with a dark line. Bolded letters are shown only for positional reference. B. Sequences from the upper strands of *sbmA* and *sbmB* are aligned and a consensus sequence is generated below. The strong consensus sequence represents bases that are present in at least 3 out of 4 of the sequences, and the weak consensus sequence represents all bases present in all 4 sites. The lower strands are also aligned below the consensus sequences. C. Sequences from the *sbmC* site are aligned and a consensus sequence generated below. The strong and weak consensus sequences from *sbmA* and *sbmB* are shown at the bottom as well.


A
sbmA
TTTTATAGGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCCCTGTTT
AAAATATCCTATGGCGATCGCCGCGACGATCGCCACGCAGGGACAAA
sbmB
TATTTTATATTAGGGGTGCTGCTAGCGGCGCGGTGTGTTTTTTT
ATAAAATATAATCCCCACGACGATCGCCGCGCCACACAAAAAAA
sbmC
AAAAAAAGCGGTGTCGGCGCGGCTACAACAACGCGCCGACACCGTTTTGTAGGGGTGGT
TTTTTTTCGCCACAGCCGCGCCGATGTTGTTGCGCGGCTGTGGCAAAACATCCCCACCA
B
sbmA upper strand
CCGCTAGGGGCGC
CTGCTAGCGGTGC
sbmB upper strand
ATATTAGGGGTGC
CTGCTAGCGGCGC
strong consensus from upper strands
CTGCTAG GGYGC
weak consensus from upper strands
$^{C}_{A}$YRYTAG$^{G}_{C}$GGYGC
sbmA lower strand
CCGCTAGCAGCGC
CCCCTAGCGGTAT
sbmB lower strand
CCGCTAGCAGCAC
CCCCTAATATAAA
C
sbmC upper strand
AAAAAAAGCGGTGTCGGCGCGGCTACAACA
sbmC lower strand
CTACAAAACGGTGTCGGCGCGTTGTTGT
sbmC consensus
A AAARCGGTGTCGGCGCG Y
$^{C}_{A}$YRYTAG$^{G}_{C}$GGYGC
CTGCTAG GGYGC

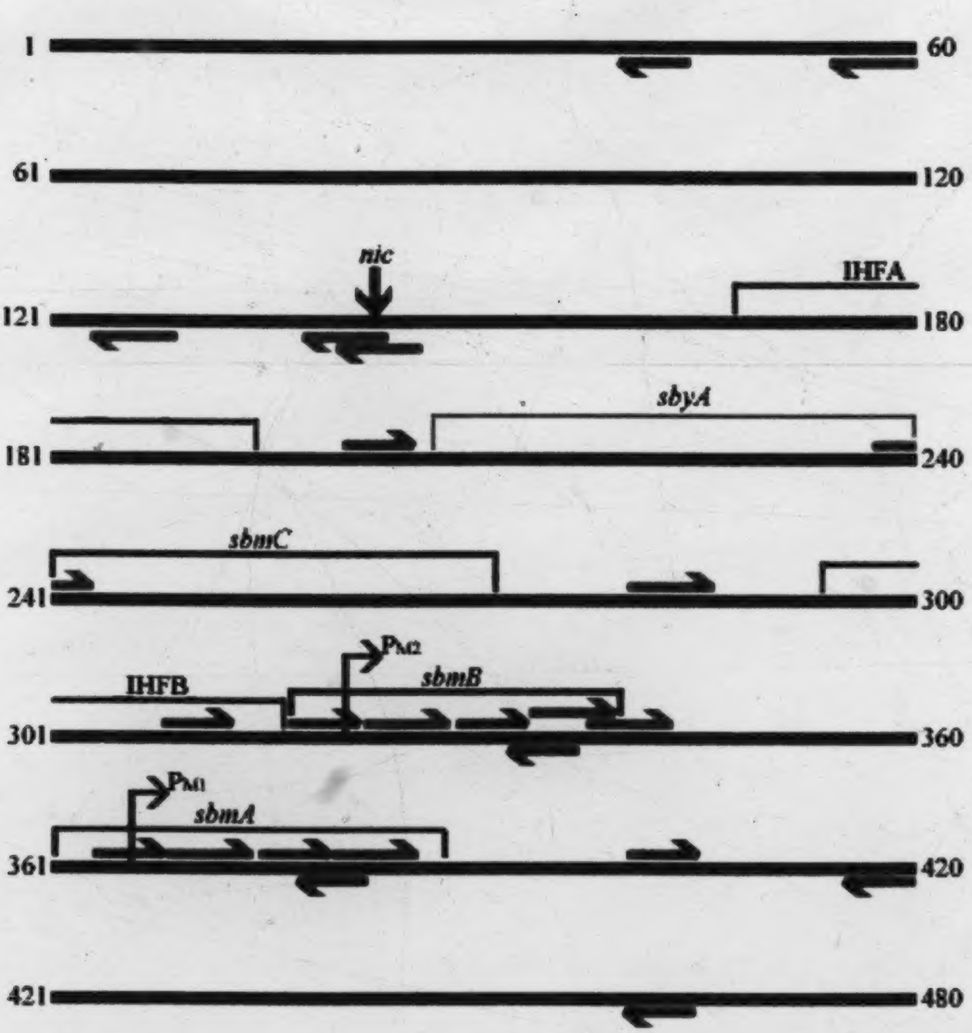
These also appear to have similarity to the consensus sequence from the upper strands, however, it should be noted that the consensus sequence has some palindromic properties centered around the CTAG motif. This may cause the appearance of sequence similarity between the upper and lower strands. The relevance of the lower strand sites needs further study in order to determine its significance in TraM binding. Figure 6.21c shows both sides of the *sbmC* site, which is an inverted repeat. A consensus sequence for the site is also generated below the sequences. Below this is the strong and weak consensus sequences for the *sbmA* and *sbmB* sites. This is to show that there are some similarities between all three of the sites, especially in the 3' end of the consensus sequence.

A motif search was done in the *oriT* region of the F plasmid using the weak consensus sequence from *sbmA* and *sbmB* (Figure 6.22). The consensus sequence was divided into the 5' and 3' ends. This was done because the 5' end of the consensus sequence was more heavily footprinted during hydroxyl radical footprinting, however, the 3' end of the consensus showed a higher degree of sequence conservation. The *nic* site is shown with a red arrow and the binding sites are shown with colored boxes. The *traM* promoters, P_{M1} and P_{M2} are also shown. The 5' end of the consensus is shown with a magenta colored arrow and the repeats are shown above and below the lines to represent positions on the upper and lower strands respectfully. The 3' end of the consensus is shown with a green arrow and repeats are shown above and below the lines to represent positions on the upper and lower strands respectfully. Statistically the 5' end of the consensus should only occur approximately every 66 base pairs. However, in the 420 bases shown this repeat is found 11 times, and is most concentrated in the last 120 bases. The 3' end of the

Figure 6.22. A diagrammatic representation of the *oriT* region of the F plasmid showing the positions of the TraM, IHF and TraY binding sites. The position of *nic* is also shown and numbering starts at the *Bgl*II site as in Frost *et al.*, 1994. Homology searches done using the first half and second half (magenta and green respectively) of the weak consensus sequence from Figure 6.21 was performed and their positions are shown above and below the DNA strand to represent upper and lower strand orientations respectively.

Weak consensus sequence: A/C YRYTAG C/G GGYGC

5' side of consensus: YRYTA 3' side of consensus: GNGGYG


1
60
61
120
nic
IHFA
121
180
sbyA
181
240
sbmC
241
300
P_{M2}
IHFB
sbmB
301
360
P_{M1}
sbmA
361
420
421
480

consensus should only be found approximately every 258 bases, however it is seen 11 times in this 420 bases. Its concentration around the *nic* site should also be noted, however, its relevance is not understood. Even though these sequences are found so frequently throughout this region, their importance is questionable since many are not found beside each other as in *sbmA* and *sbmB*.

Discussion

The DNA binding properties of the newly purified and biochemically characterized TraM were examined. This was done using a variety of templates in order to characterize the protein as thoroughly as possible. Using different preparations of TraM, crude and homogeneous, DNA binding properties had been previously analyzed using EMSA (Penfold, 1995) and DNase I footprinting assays (Di Laurenzio, 1992). Results presented here agree with those previously published which showed that *sbmA* is the strongest binding site for TraM, and *sbmC* is the weakest. These experiments also showed that *sbmABC*, representing all of the binding sites on one fragment, bound to TraM with a 5 fold higher affinity than *sbmAB*. However, in addition to these previously known characteristics, a more thorough analysis provided a better idea of how TraM binds to its binding sites.

An interesting phenomenon involved the binding patterns of *sbmA*, *sbmAB*, and *sbmABC* versus *sbmC* and a half *sbmA* site. The binding pattern of the first three gave defined species suggesting the formation of distinct molecules which are composed of specified amounts of protein and DNA. However, binding to the latter two binding sites gave species which increased in size as more TraM was added. This suggests that as more protein is present in the system, the protein complex bound to the DNA also increases in size. The smearing of these species during EMSA (Figure 6.4) also suggests that these complexes are not stable and dissociate while the gel is running.

Cooperative binding of TraM to its binding sites had been previously hypothesized (Penfold, 1995) and was therefore tested. EMSA experiments performed here also suggested TraM binds cooperatively to some of its binding sites. To analyze this, four quantitative measurements were made. The first used Hill plots to determine cooperative binding. *sbmA*, *sbmAB*, *sbmABC* all fulfilled the requirements needed to conclude that these sites are bound cooperatively. The second was measuring the breadth of the binding curve taken from 10% to 90% bound DNA. Cases where this increase occurs in less than 1.81 log units of protein are assumed to be cooperative. This was found to be the case for *sbmA*, *sbmAB*, *sbmABC*. Another analysis was performed using τ, which measures the increase of the intermediate bound complexes during EMSA experiments. Buildup of these intermediate bands suggests that cooperative binding does not occur. Since only *sbmAB* and *sbmABC* gave more than one species during EMSA only these fragments could be measured using this method. Both fragments showed that cooperativity occurred during protein binding. The last analysis was performed using ω, which determines cooperativity using the binding constants of the sites alone and when present on the same DNA fragment. Since the binding constant for half of *sbmA* had been determined, cooperativity of *sbmA* could be calculated and was found to be present. This may result from the ability of TraM proteins to assist each other in the binding of half *sbmA* sites. Cooperativity was also found for *sbmAB* using this technique.

During DNase I footprinting assays TraM was found to only protect the non-transferred (upper) strand of the *oriT* (Di Laurenzio *et al.*, 1992). EMSA of oligonucleotides showed that TraM does not bind the lower strand but does bind the upper strand of the *sbmA* site.

However, closer inspection of the gels using the upper strand showed another species above the unbound single-stranded oligonucleotide that was also bound. The possibility that this species was a dimer was complicated by the fact that the lower strand, which was perfectly complementary to the upper, did not have this species during EMSA. However, during dimerization, the upper strand may create more stable pairing than the lower. For example, if the two oligonucleotides were paired in the same manner a G-T pairing in the upper oligonucleotide dimer would be matched by a C-A pair in the lower dimer. Since the G-T pair is more stable than the A-C pair, situations like this could be responsible for seeing a dimer with the upper strand but not with the complementary lower strand. To test this, all combinations of dimers were examined to determine whether more stable pairing could be seen with the upper strand. However, none were found. Another possibility is that this upper species represents a hairpin of this oligonucleotide. An oligonucleotide was synthesized to destabilize this type of structure and showed less of this upper species during EMSA, however, still bound TraM with the same affinity. This suggested that the upper species may be a hairpin, however, was not required for TraM binding. Changing the spacing between the half sites in *sbmA* virtually eliminated TraM binding suggesting that the distance between the half sites is important. Binding to a single-stranded half *sbmA* site was also of interest, however, it was found that TraM would not bind to an oligonucleotide of 16 bases. When this oligonucleotide was extended 5' and 3' by 10 bases binding was seen to the single-stranded half site. This confirmed the previous data that TraM does indeed bind single-stranded DNA.

TraM was also seen to bend its binding sites to a small degree once bound. *sbmA* was the most bent of all of the sites while *sbmC* was the least bent. Once TraM is bound this bending may aid in the binding of more TraM to other sites and participate in the overall quaternary structure of the *oriT* region. Bending of *sbmAB* and *sbmABC* was also observed, however, this assay cannot distinguish wrapping of the DNA which may come from the linking of binding sites on one DNA molecule. Therefore, results for fragments having more than one binding site are questionable.

During competition assays all of the sites, except *sbmC*, showed a low level of dissociation. Even at 1000-fold excess of specific competitor very little dissociation of TraM from its sites was seen, suggesting a very stable complex is formed upon binding. Binding to *sbmC* was not as strong, and dissociation was seen when competitor reached a 100-fold excess. Small amounts of dissociation were seen in the upper complexes of *sbmAB* and *sbmABC* at high levels of competitor. However, these never dissociated from the middle species to give unbound DNA. This suggests that the upper species may be somewhat more sensitive to dissociation, while the middle species is very stable.

Hydroxyl radical footprinting was also performed on TraM bound to *sbmA* and *sbmC* in order to define the protected bases in the previously published DNase I footprint (Di Laurenzio *et al.*, 1992). Footprinting of *sbmA* showed that TraM bound both the upper and lower strand which was not expected and is not easily explainable since TraM did not bind the lower strand during EMSA of single-stranded oligonucleotides. However, it should be noted that this assay only detects the protection of specific bases and does not necessarily mean these bases are required for binding. These bases on the lower strand may not be required for protein binding but are protected once the protein is bound. The footprinting of *sbmA* also showed that TraM bound to the bases upstream of the CTAG sequence in the *sbmA* consensus sequence (Figure 6.20). This was unexpected, however, the footprint was the same for the upper and lower strands. Whether the protection is due to the spatial position of these bases or these are the bases recognized by TraM will require further analysis. Modeling TraM onto this site suggests that 2 major grooves are bound by the protein, and that the protein folds around 2 faces of the DNA in this major groove. One protein molecule may occupy each of the grooves and form a dimer through the intermolecular interactions of TraM. Tetramerization may result from two TraM dimers, each bound to *sbmA* and *sbmB* in a similar fashion. Footprinting of *sbmC* showed that TraM also bound the top and bottom strands of this site. This site seemed to be protected in a more symmetrical fashion centered around the ACA in the center of the DNase I footprint. Footprints suggest that TraM is bound to one side of the DNA on one end of the site, and to the opposite side on the other end of the site. Modeling TraM onto the *sbmC* site may be possible as two dimers, one bound to the major grooves on the upper side at one end of the site, and another dimer bound to the major grooves on the lower side at the other end of the site. Tetramerization of the protein may bend and wrap the DNA around it so that the bases in the center of the site are protected non-specifically.

Comparing the sequences from *sbmA* and *sbmB* produced a consensus sequence, however, only the 5' end of this was footprinted during hydroxyl radical experiments.

Comparison of this consensus sequence to the consensus sequence for *sbmC* also showed the most homology at the 3' end. Two possibilities can explain this paradox: First, the sequences recognized by TraM are not protected upon binding. Second, sequence gazing for TraM consensus sequences may not be as informative as for other proteins. Mutagenesis of these sites to determine if the 5' or 3' end of the consensus sequence is important and will need to be done in order to answer these questions. Scanning the *oriT* region for the consensus sequences showed that these sites are present quite often. However, mutagenesis to determine the important bases in the binding sites will determine whether or not these sites are of importance.

Chapter 7

General Discussion

General Discussion

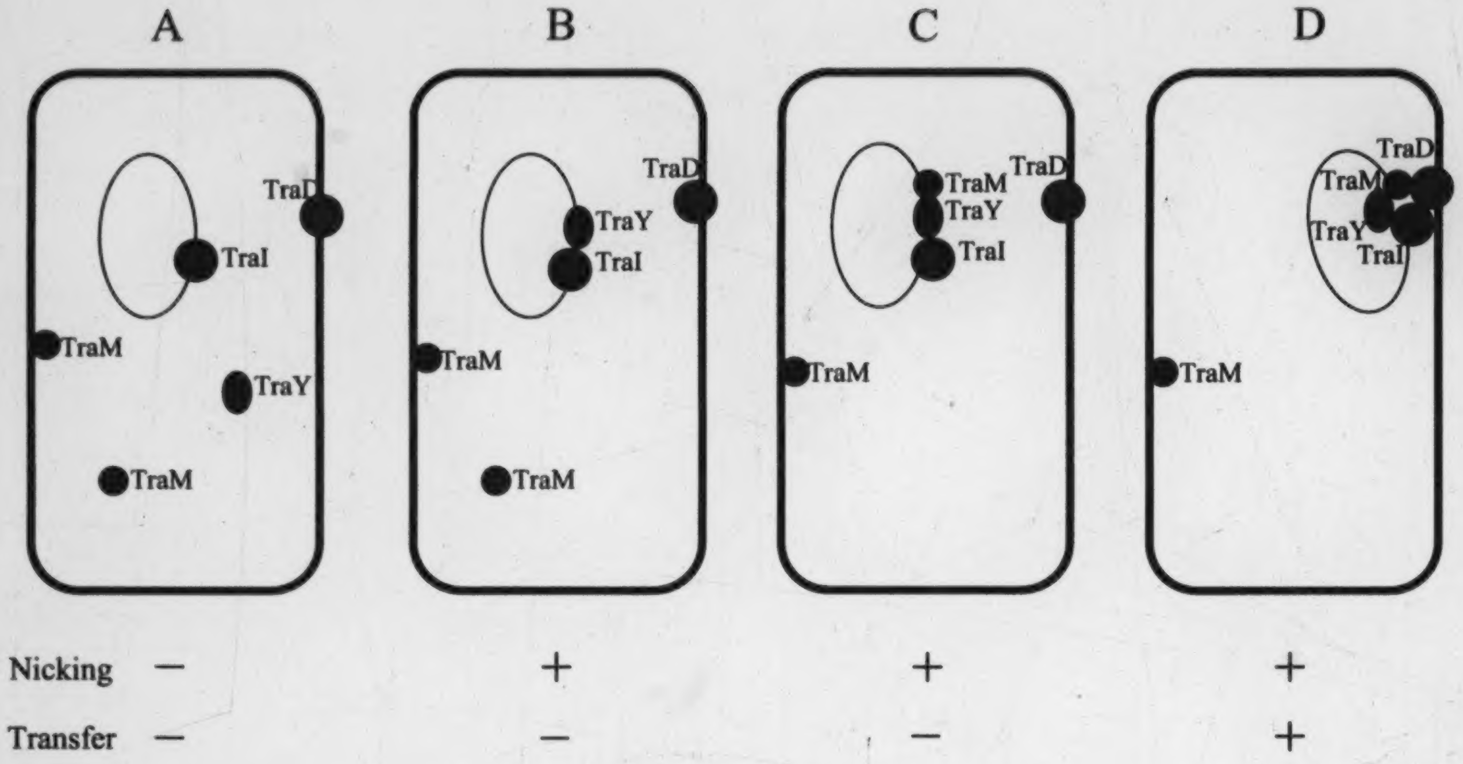
TraM is essential for successful transfer of the F plasmid (Achtman *et al.*, 1971). It is thought to participate in the mechanism for signaling that a stable mating pair has formed and that transfer may occur (Willetts and Wilkins, 1984). This was based on the observations that TraM is not required for pilus formation (Achtman *et al.*, 1972) or nicking at the *nic* site (Everett and Willetts, 1980), but is required for DNA transfer and subsequent steps. TraM was originally localized to the inner membrane (Achtman *et al.*, 1979; Thompson and Taylor, 1982), however, was later localized to the cytoplasm with small amounts still found in the membrane (Di Laurenzio *et al.*, 1992). The mutant phenotypes and localization pattern make TraM a candidate for a role in the signal mechanism. TraM from the F-like R1 plasmid is 78% identical to F TraM, with most of the differences occurring in the amino-terminus of the protein. R1 TraM is required for nicking *in vivo* (Kupelwieser *et al.*, 1998), and for expression of the pilus subunits (Polzleitner *et al.*, 1997). Even though the two proteins share a high level of identity, these differences in function suggest that the two proteins perform different functions in the transfer of their respective plasmids.

The F-like plasmid R100-1 shares a great deal of similarity to the F plasmid *oriT* region at the levels of sequence and binding site arrangement. Experiments which mixed *oriT* DNA binding sites of these two plasmids demonstrated TraM's allele-specific nature, and showed that TraM must be bound in *cis* to *nic* in order to perform its function in DNA transfer. These experiments also showed that TraM's allele-specific nature is based on

DNA binding alone, and once bound either F or R100-1 TraM can function with the relaxosome and transfer apparatus. Quantitative levels of cleavage by TraI were also only seen when TraY was able to bind in *cis* to *nic*. These results support the previous observations of Howard *et al.* (1995) which showed that TraY was required in order to allow nicking by TraI at the *nic* site. This data suggests that an unknown quaternary structure is formed when the TraI and TraY proteins are bound to the *oriT* region, allowing nicking. The subsequent binding of TraM then allows transfer of the plasmid DNA to occur (Figure 7.1).

TraI is suggested to interact with the inner membrane through TraD (Dash *et al.*, 1992). Recently TraG from RP4 (the F TraD homologue) was shown to directly interact with the relaxase (Mob) from the mobilizable plasmid pBHR1 (Szpirer *et al.*, 2000). The cytoplasmic localization of TraM (Di Laurenzio *et al.*, 1992) and the demonstrated interaction between F TraM and the inner membrane protein TraD (Disque-Kochem and Dreiseikelmann, 1997) suggest that TraM may anchor the plasmid DNA to the inner membrane. However, cell fractionation experiments suggest that TraM is only peripherally bound to the inner membrane and that this localization is not affected by mutations in TraD. Gordon *et al.* (1997) showed that the F plasmid is localized to the one quarter, one half, and three quarter positions of the cell. It is therefore possible that transfer occurs only at selected positions within the cell, and that the specific interaction between TraM and TraD may only occur at these positions. When comparing TraD mutants to wild-type cells this limited interaction may not be frequent enough to detect TraM in membrane fractions using Western analysis. This interaction may also be totally

Figure 7.1. Conjugation steps prior to transfer as mediated by TraM, TraY, TraI, and TraD. Diagrams are shown to represent plasmids in various *tra* backgrounds. The order in which they are shown is not meant to signify the order in which binding occurs *in vivo*. Circles may represent more than one molecule of each protein. The ability of each cell to nick and transfer is also designated below the figure. A. Diagram representing pRF206 in an F background. This plasmid contains a R100-1 *oriT* region with F *nic*. Nicking of pRF206 is not visible in nicking assays, however, some is presumably present since this plasmid is transferred. B. Diagram representing pOX38-*traM*K3 and pRF105 (supplying R100-1 TraM). F TraY and TraI are able to bind to pOX38-*traM*K3 and nicking is observed, however, R100-1 TraM is unable to bind resulting in no transfer of pOX38-*traM*K3. C. A proposed timepoint in the F plasmid cycle where all *tra* proteins are present, however, due to plasmid replication or cell division, the plasmid is not in the correct position and transfer is unable to occur. D. Diagram depicting the F plasmid just prior to transfer where all *tra* gene products are present and functioning allowing nicking and transfer to be detected.


A
B
C
D
TraD
TraI
TraM
TraY
TraM
TraD
TraY
TraI
TraM
TraM
TraM
TraY
TraI
TraD
TraM
TraD
TraM
TraY
TraI
TraM
Nicking − + + +
Transfer − − − +

separate from the non-specific peripheral membrane association seen with TraM. These possibilities are shown in Figure 7.1. The positioning of the plasmid to the one quarter, one half, and three quarter positions in the cell are not suggested to be due to the position of TraD or the pilus, but transfer may only be able to occur when TraM and/or TraI is able to interact with TraD at the inner membrane.

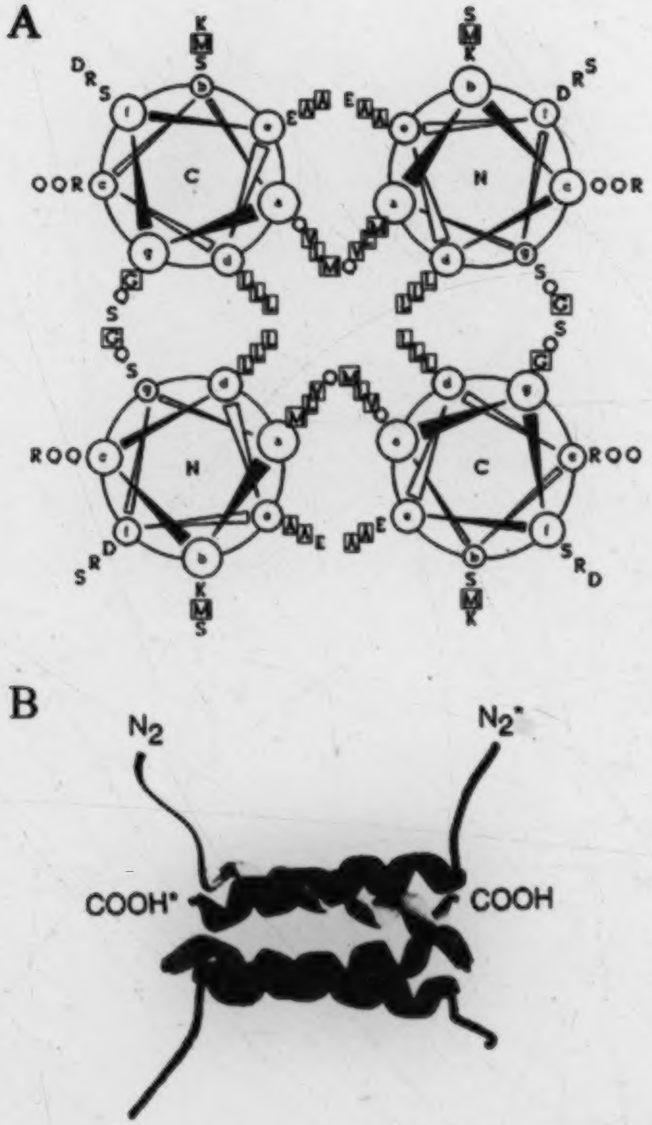
In the present work TraM was overexpressed, purified and shown to have a molecular weight of 14 376 Da. Amino acid analysis also showed that the amino-terminal methionine is cleaved from the protein. Analytical ultracentrifugation and size exclusion chromatography showed native molecular weights of between 56 000 and 66 000 Da, suggesting that TraM exists as a tetramer in solution. This confirms the results of sucrose gradient centrifugation using purified protein (Di Laurenzio, 1992), and size exclusion chromatography using crude cell extracts (Penfold, 1995). Chemical crosslinking using a variety of reagents suggests dimers and tetramers of TraM in solution. Determining whether more than one type of dimer was formed was not possible, however, results suggest that there are two interacting domains in TraM: one which forms dimers at low protein and/or crosslinker concentrations, and one which forms tetramers at higher concentrations. The domain involved in dimer fromation appears to be strong since some dimer remained after denaturation during SDS PAGE. Crosslinking profiles of TraM did not change upon binding to its binding sites, which has been observed with other proteins such as T4 RegA (Phillips *et al.*, 1996) and with the retinoid X receptor (Kersten *et al.*, 1995). These proteins appear to undergo a conformational change when they bind to DNA, changing the crosslinking profile at defined protein concentrations.

The yeast two-hybrid system was used to further characterize the domains of TraM responsible for dimerization and tetramerization. Analysis of full length and deletion mutants suggested that there are two regions of interactions in TraM. The first is the central region (amino acids 25 to 108) which may interact with the same region of another TraM subunit. The second is the carboxyl-terminal region (amino acids 108-127) which may also interact with the same region of another TraM subunit. The central region is predicted to be alpha helical and amino acids 25 to 85 are amphipathic. This has been suggested to be characteristic of regions involved in protein:protein interactions. R1 TraM is 83% identical to F TraM in this central region, and dimers have been predicted to form when the first 80 amino acids of the protein are analyzed (G. Koraimann, personal communication). This suggests that this region may be the dimerization domain of TraM. The Lac repressor also has a dimerization domain in the central part of the protein, and like TraM, contains a DNA binding region in the amino-terminus (Friedman *et al.*, 1995).

The carboxyl-terminus of TraM also appeared to be involved in protein:protein interaction, and deletion analysis of *traM* using yeast two-hybrid analysis suggests that the carboxyl-terminus most likely interacts with the same region of another TraM subunit. This region is also predicted to be alpha helical and amino acids 101 to 127 are amphipathic. This is similar to the Lac repressor, which has an 18 amino acid amphipathic alpha helix joined to the body of the protein by a 7 amino acid extended coil (Friedman *et al.*, 1995). This region of the Lac repressor has been designated the tetramerization or oligomerization domain and forms a four-helix bundle using the carboxyl-terminal regions of four monomers. Unlike TraM, this domain contains leucine heptad repeats, however, the similar amphipathic nature is noteworthy. Figure 7.2a shows the four-helix bundle of the Lac repressor from the end, demonstrating its amphipathic nature, and Figure 7.2b shows a ribbon diagram of the bundle. R1 TraM is 89% identical to F TraM in the carboxyl-terminal 27 amino acids and is predicted to have a carboxyl-terminal alpha helix bound to the body of the protein by a loop of 19 amino acids (Verdino *et al.*, 1999). This region of R1 TraM has also been postulated to form a four-helix bundle (G. Koraimann, personal communication) and to be the tetramerization domain. Similarities between these systems suggest that the carboxyl-terminal region of TraM may be the tetramerization domain of the protein.

Mutagenesis of amino acid residues in the helices which appear to be structurally important could be used to determine the functions of the amphipathic helices. For example, acidic residues could be replaced with basic ones, or hydrophobic residues could be replaced with charged ones. Similar experiments were done with the Lac repressor where an apolar substitution (tyrosine for aspartic acid) reduced monomer-monomer stability (Nichols and Matthews, 1997). To determine whether the helices are involved in inter- or intra-molecular interactions, a screen for second site suppressors of mutations could be performed. This can be accomplished using non-specific PCR (Nichols and Matthews, 1997) where misincorporation of bases occurs at high frequencies. Analysis of TraM with the eight amino acid carboxyl-terminal deletion (Penfold, 1995; Frost *et al.*, 1997) is also important to determine why this protein is non-

Figure 7.2. The four-helix bundle structure of the Lac repressor. A. The bundle shown from the end, showing amino acids arranged in a helical wheel diagram. Hydrophobic amino acids are shown with a box around them. Modified from Alberti *et al.* (1993). B. A ribbon diagram of the four-helix bundle showing anti-parallel arrangement. The linker arms connecting the carboxyl-terminal alpha helix to the body of the protein are also shown. Modified from Friedman *et al.* (1995).


A
C
N
N
C
OOR
ROO
B
N2
N2*
COOH*
COOH

functional. Removal of these last eight amino acids could weaken the ability of TraM to tetramerize enough so that it can no longer cooperatively bind DNA, however, is still seen as a tetramer in solution.

Analysis of the protein:DNA interactions of TraM was performed in order to understand how each of its sites are bound, and what the effects of binding are on the DNA. Characterizing the binding of F TraM to each of its binding sites alone and in combination showed that TraM had the highest affinity for fragments containing all three sites (*sbmABC*) and the lowest affinity for *sbmC* alone. Binding patterns and affinities also varied between each of the binding sites. Binding to *sbmC* showed smearing and complexes increased in molecular weight as more protein was added to the reaction. This suggested that binding to *sbmC* did not form specific complexes of protein and DNA. Binding to this site did not demonstrate cooperativity, and TraM did not remain bound when *sbmA* or *sbmABC* was added in competition assays. Analysis of TraM bound to this site also demonstrated very little bending of the DNA template. Sizing of the bound DNA complex suggested that a tetramer may be bound at 500 nM TraM.

Binding of TraM to *sbmA* on the other hand demonstrated stable binding and formation of a specific protein:DNA complex. This complex formed larger complexes at levels indicative of nonspecific binding. TraM:*sbmA* complexes were not dissociated even when a 1000-fold excess of *sbmA* or *sbmABC* competitor was added. Sizing of the bound TraM molecule suggested that first one dimer binds, followed by another to form a tetramer. This initial binding step was not seen during EMSA and may be due to the high levels of cooperative binding at this site. Using the pBEND2 vector and applicable restriction enzymes, TraM is suggested to bend *sbmA* to an angle of approximately 50°. This bending may assist the subsequent binding of protein to other sites on the DNA.

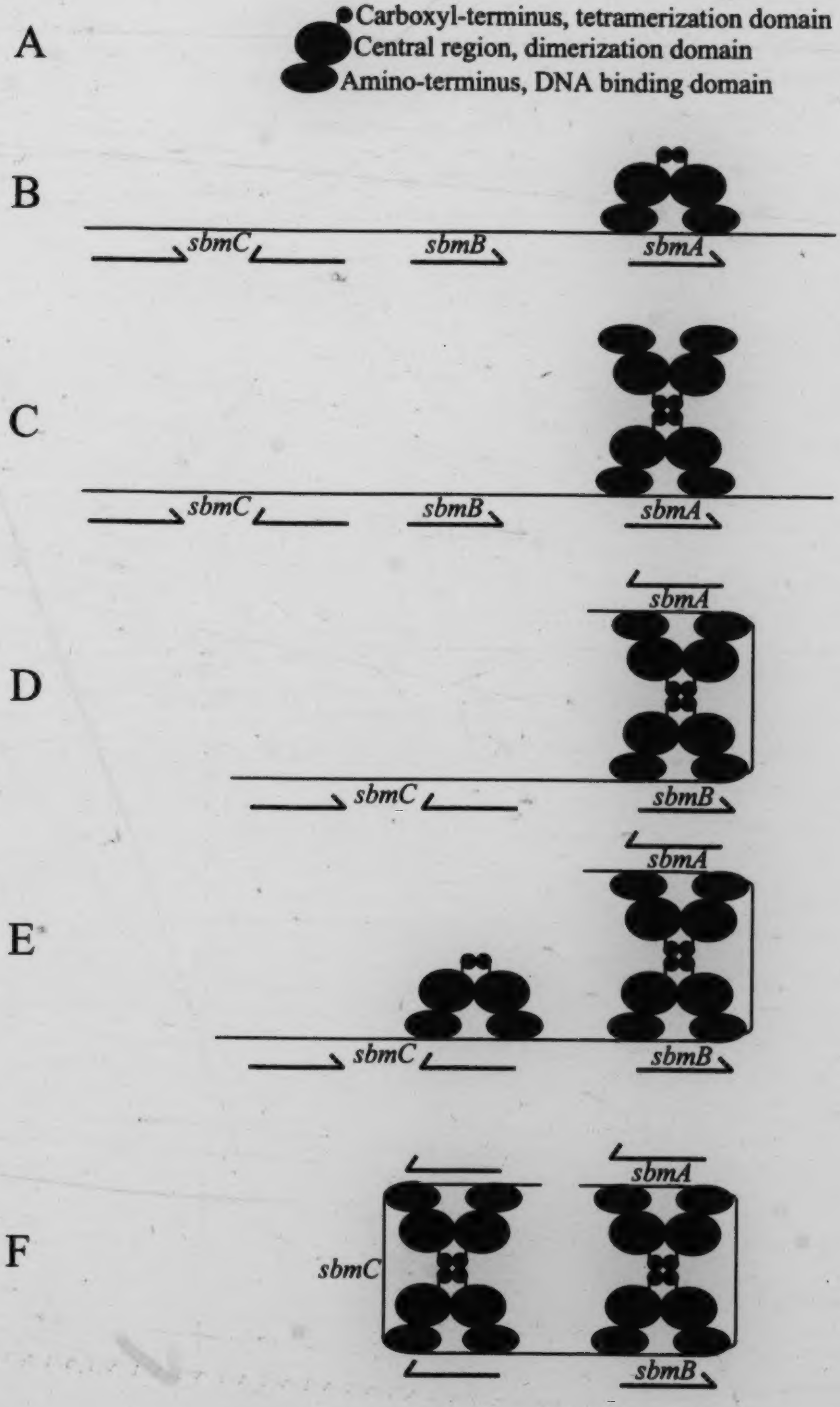
Binding of TraM to *sbmAB* produced two stable species during EMSA and sizing of the final complex suggested that a tetramer was bound to the two sites. Data suggested cooperative binding to these sites and during competition assays the upper complex was lost, but the first protein:DNA complex seemed much more stable and was not dissociated. Experiments to identify these two complexes proved unsuccessful. Binding to all of the TraM sites in *oriT* (*sbmABC*) also produced two stable species during EMSA experiments. Sizing of the upper bound complex suggested that a hexamer of TraM was bound to these sites. Cooperative binding was shown for these three sites which also showed the highest affinity for TraM of all of the binding site combinations. As with *sbmAB*, competition experiments using *sbmA* and *sbmABC* showed a loss of the upper band, but little dissociation of the middle complex to unbound DNA was seen.

The single-stranded binding property of TraM was demonstrated using full-length *sbmA* sites and half *sbmA* sites. The spacing between the half sites was shown to be critical, suggesting that some type of protein interaction occurs between the TraM molecules bound to each half site. Interestingly, TraM bound only to the upper strand of *sbmA*. However, this oligonucleotide also showed some peculiar properties. This oligonucleotide ran as 2 bands during EMSA in the absence of protein suggesting that dimerization or hairpinning may be occurring. Decreasing amounts of the intermediate band were observed upon modification of the DNA sequence to eliminate proposed hairpin structures, while maintaining TraM consensus binding sequences. The observation that one strand forms secondary structures and the other does not may be explained by specific DNA sequence effects, where one strand forms more stable base pairs than the other when complementary structures are formed.

Hydroxyl-radical footprinting of TraM bound to *sbmA* and *sbmC* was also performed to obtain a better understanding of which bases are protected in the previously published DNase I footprints (Di Laurenzio, 1992). Surprisingly, bases on the upper and lower strands of *sbmA* were protected. However, determining which are protected due to base recognition versus non-specific base protection could not be determined using this assay. Future mutational analysis of the most conserved bases may clarify which bases are required for TraM binding. The bases before the conserved CTAG were most heavily protected on the upper and lower strands of the TraM binding sites. However, the sequences after the CTAG share the most homology. Footprinting data suggested that TraM binds to the *sbmA* site in two major grooves, each by one protein molecule. Footprinting of *sbmC* showed a high degree of symmetry centered at the middle of the binding site, with bases protected on the upper and lower strands. A weak consensus was observed between the protected bases of *sbmA* and *sbmC*, however, mutational analysis of the conserved sites in *sbmA* may provide some insight into how TraM can bind two different binding sites. One proposed model for the binding of TraM to *sbmC* suggests that two major grooves are protected on each side of the site by a dimer. Tetramerization of the dimers may twist and wrap the DNA around the multimer so that bases in the center of the site are subsequently protected non-specifically by the side of the tetramer.

A clear model for the binding of F TraM to the *oriT* region of the F plasmid does not yet exist. Defining specific interactions is not possible without intensive mutational analysis and crystal structures. However, experiments performed to this point give an indication of how binding may occur. EMSA results suggests that *sbmA* has the highest affinity for TraM, thus this is probably the site which is initially bound by the protein. Sizing of bound TraM to its sites suggests that a dimer binds to *sbmA*, quickly followed by tetramerization. Dimer binding to *sbmA* could be similar to the binding of the Lac repressor to its sites (Lewis *et al.*, 1996) where a dimer is bound to a single site having both tetramerization domains positioned away from the DNA (Figure 7.3B). This explains the importance of spacing between the *sbmA* half sites to allow proper interaction between the monomers at the dimerization interface. Since the dimer:dimer interaction of TraM is strong enough to allow tetramers in solution, *sbmB* is most likely bound after tetramerization (Figure 7.3C and D) as is the case for the Lac repressor which binds two operator sequences as a tetramer (Kramer *et al.*, 1987). However, these two steps (tetramerization and binding *sbmB*) may not occur in this order since it is possible that another dimer binds to *sbmB* and tetramerization of the two dimers wraps the DNA around the protein. An example of the latter is the cooperative binding of *E. coli* LexA where monomers bind to the binding site in sequential fashion instead of dimerizing in solution (Kim and Little, 1992). Cooperative binding of LexA is achieved through protein:protein contacts between the bound monomers. Binding of another dimer to

Figure 7.3. Model suggesting how TraM binds to its binding sites in the *oriT* region of the F plasmid. DNA sequence lengths and protein sizes are not shown to scale. Blue arrows show the direct and indirect repeats shown in blue in Table 6.1.


A
Carboxyl-terminus, tetramerization domain
Central region, dimerization domain
Amino-terminus, DNA binding domain
B
sbmC
sbmB
sbmA
C
sbmC
sbmB
sbmA
D
sbmA
sbmC
sbmB
E
sbmA
sbmC
sbmB
F
sbmA
sbmC
sbmB

sbmC then follows and is in some way stabilized by the tetramer bound to *sbmAB* (Figure 7.3E). Binding of TraM to *sbmC* may also aid in the binding of TraM to *sbmAB* since binding to all three sites together is cooperative in comparison to *sbmAB* alone. During previous analysis, TraM had shown some aggregation characteristics (Di Laurenzio, 1992) which may explain the ability of protein molecules bound at *sbmAB* and *sbmC* to interact. At higher protein concentrations another dimer may bind to *sbmC*, forming another tetramer at this binding site (Figure 7.3F). DNA at this site may be twisted in order to allow DNA binding of each of the dimers of the second tetramer, however, the *in vivo* significance of this form is not understood. Testing of this model requires the definition of the protein:protein and protein:DNA interacting domains of TraM. Determination of specific sequences bound by the protein will also allow for a more defined model of TraM bound to the DNA.

The role of TraM is still not clear, however, it is clear that DNA binding must occur in *cis* to the origin being transferred. TraM's interaction with the inner membrane protein TraD suggests that it may anchor the transferred plasmid to the base of the pilus. However, this may only occur just prior to, and during DNA transfer. TraM has been shown to affect the superhelical density of DNA (Di Laurenzio, unpublished results), and the same has been suggested for R1 TraM (E. Zechner, unpublished results). Since TraM is not needed for nicking in the F plasmid, TraM may change the superhelical density in the *oriT* region upon binding and cause further unwinding at the *nic* site. Ironically, this theory is very similar to one first suggested by Everett and Willetts in 1980 and should be examined using supercoiled substrates during EMSA and during potassium permanganate footprinting of the *nic* site. The former would determine if supercoiling has an effect on TraM binding, and the latter would determine if TraM has an effect on the amount of unwinding at *nic*. Unwinding may allow another step to occur such as the switching of TraI from a relaxase to a helicase, which would then begin to separate the transferred and non-transferred strands. TraM could still be part of the "signal" informing the donor cell that a recipient is ready (as originally suggested by Willetts and Wilkins, 1984). TraM's function in the signal could be postulated to occur at any point in any of its postulated functions (illustrated in Figure 7.1). These could include binding to DNA, changing the conformation of the DNA by saturation of its binding sites, and bringing the plasmid to the pilus through TraM:TraD interactions. The determination of TraM's specific role in the transfer process will require further study, including mutations of the protein and its binding site to separate these putative functions.

Chapter 8

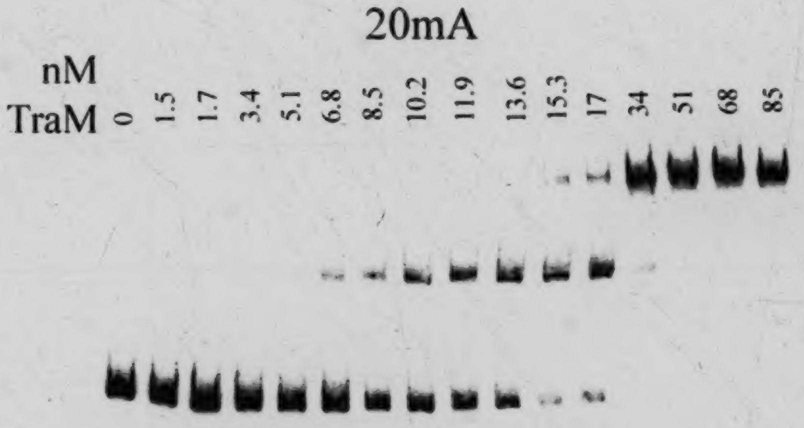
Appendix

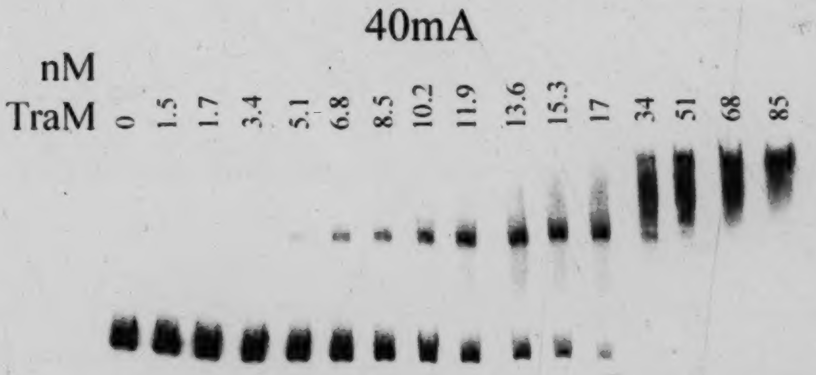
Appendix

Creating an epitope tagged TraM fusion. For simple purification of TraM protein, a histidine tag was fused to the amino-terminus of *traM*. Placement of six histidine residues on a protein allows purification of the fused protein using a nickel-NTA affinity column (Qiagen). pRF400 (wild type TraM), pRF401 (carboxyl-terminal eight amino acid deletion), and pRF402 (point mutation changing isoleucine at position 109 to threonine) were created by placing these genes into pQE40 (Qiagen), described in "Materials and Methods". Transfer of pOX38-Km and pOX38-*traM*K3 was assayed in the presence of this protein as described in the Mobilization efficiency assays in "Materials and Methods". Expression of the TraM fusions was induced by the addition of IPTG to 1 mM 1.5 hours prior to mixing donor and recipient cells. The transfer of pOX38-Km decreased approximately 100-fold in the presence of pRF400 and pRF401 compared to a control plasmid (pUC19), and transfer decreased approximately 10-fold in the presence of pRF402. Transfer of pOX38-*traM*K3 was not facilitated in the presence of any of the TraM fusions. Negative effects of IPTG on transfer were ruled out using a control plasmid (pNY300) which facilitated the transfer of pOX38-*traM*K3. This suggests that fusing six histidine residues to the amino-terminus of TraM eliminates its function. This may be because the DNA binding domain (located in the amino-terminus) is physically blocked and is unable to bind DNA properly. The decrease in the transfer of pOX38-Km (which supplies wild-type TraM) suggests that wild-type function is poisoned in the presence of the fusion proteins. This may be explained by formation of inactive hetero tetramers. This further suggests that the multimerization domains of the fused TraM proteins are not altered and requires further investigation.

Effects of running conditions on EMSA. Various conditions were altered to determine their effect on EMSA of the TraM binding sites. Changing the temperature at which the gel was run from 4$^{\text{o}}$C to room temperature caused more smearing and a decrease in the affinity of TraM for its binding sites. This was presumably due to an increase in gel temperature, which resulted in a destabilization of the complexes. Altering the buffer composition in which the gels are run from Tris-Borate to Tris-Acetate, Tris-Glycine, and Tris alone has been shown to affect the affinity of proteins for their binding sites (Urh *et al*., 1995; M. Filutowicz personal communication). To determine if this was the case for TraM electrophoresis was performed in various buffers. Except for an altered mobility of the complex, no effects on affinity of TraM for its sites was observed. This was in contrast to previous preparations of TraM which were significantly affected by buffer composition. These previous preparations were also affected by the presence of EDTA in the buffer. This was presumably due to the different ionic characteristics of the various buffers which could directly affect DNA binding or affect DNA binding by altering the multimerization of TraM. The effect of current (20 to 40 mA) on newer protein preparations was analyzed at 4$^{\text{o}}$C. No effects were seen on *sbmA*, *sbmC*, or the intermediate complexes in *sbmAB* and *sbmABC*. However, the upper complexes in *sbmAB* and *sbmABC* were severely affected (Figure 8.1 shows *sbmAB*. Identical patterns were observed for *sbmABC*). Gels run at 30 mA gave the same patterns as those run at 20 mA. It is possible that the increase in current causes an increase in temperature which

Figure 8.1. Effects of current on EMSA of *sbmAB*. Binding sites were cloned into pBEND2 and bound by the indicated amounts of TraM shown in nM. Each lane contains 0.3 fmol of DNA which was produced by PCR using end-labeled primers. The current at which each gel was run is indicated above each gel.


20mA
nM
TraM
0
1.5
1.7
3.4
5.1
6.8
8.5
10.2
11.9
13.6
15.3
17
34
51
68
85


40mA
nM
TraM
0
1.5
1.7
3.4
5.1
6.8
8.5
10.2
11.9
13.6
15.3
17
34
51
68
85

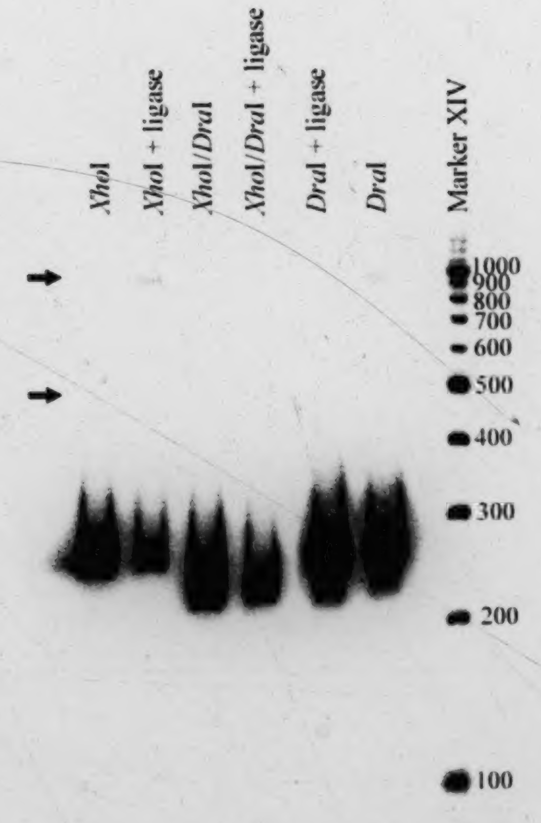
then destablizes the upper complex during the run resulting in a smeared band in the gel. Since the intermediate complex is not affected by current it suggests that it is more stable than the upper complex. The implications of these observations are unknown since the identity of the two complexes is unknown.

***sbmAB* circularization experiments.** In the model presented in Figure 7.3 TraM binds as a tetramer to both binding sites in the *sbmAB* fragment. This would be accomplished by looping the DNA around the tetramer so that the DNA binding sites are on either side. An assay was developed in order to determine if TraM was looping the *sbmAB* fragment under the assumption that bending the DNA would bring the ends of the DNA fragment closer together. Adding T4 DNA ligase to this reaction would then result in more frequent circularization of the fragment if TraM was bound, looping the DNA to cause a high local concentration of compatible DNA ends. pRF930 was used as template for PCR using primers RFE16 and RFE17 which contained 5 μL (50 μCi) of [α^{32}P]dCTP (Amersham). DNA was isolated from a 1.5% agarose gel as in EMSA experiments. One fifth of the PCR reaction was digested with *Xho*I, *Dra*I, or both *Xho*I and *Dra*I for 8 hours. Reactions were dried down to 15 μL and the 260 bp, 260 bp, and 244 bp (respectively) fragments were purified from an 8% acrylamide gel. DNA was then quantitated in the same manner as the DNA in EMSA reactions. Approximately 6 fM of DNA (approximately 10 000 cpm) was used in a 15 μL binding reaction using Ligase buffer and TraM for 15 minutes. 1 μL of T4 DNA ligase was added to the reaction and incubated at 37°C for 30 minutes. Reactions were phenol-extracted and ethanol-precipitated and run on 8% TB gels or 8% denaturing polyacrylamide gels containing 8

M urea. Gels were exposed to a Phosphor Screen (Molecular Dynamics) overnight (Figure 8.2). As a size standard, 10 μL of end-labelled Marker XIV (Roche) was dephosphorylating using Alkaline Phosphatase (Roche). DNA marker was phenol-extracted, ethanol-precipitated, resuspended in 6 μL Milli-Q® water, end-labelled using Polynucleotide Kinase and 4 μL (40 μCi) of [γ^{32}P]ATP and purified using a Nuc-Trap Probe Purification column (Stratagene). Circularized DNA fragments were expected to have decreased mobility in comparison to the monomer DNA fragments and different mobility than dimer or tetramers of the DNA fragment. These circles were also only expected to be seen only when TraM was present in the reactions. No bands of this type were seen and extra bands could be attributed to dimers or tetramers (Figure 8.2, lower and upper arrows, respectively) or were not reproducible. A possible explanation is that if the model presented in Figure 7.3 is correct the DNA ends may not be flexible enough to interact with each other due to a physical restraint imposed by the bound TraM.

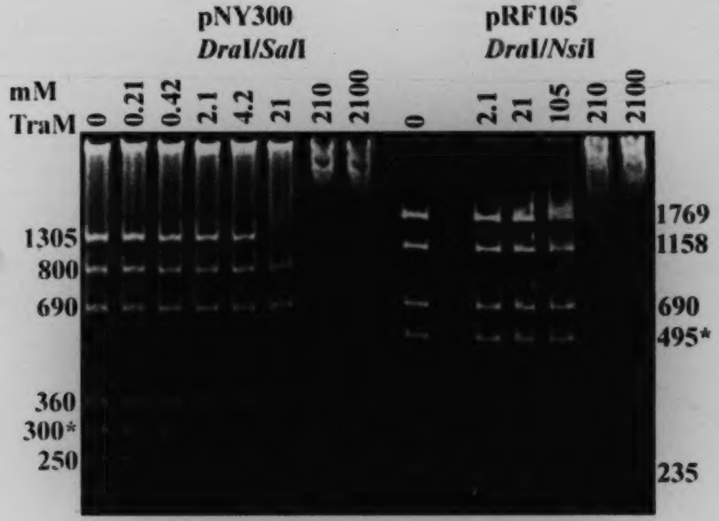
Binding of F TraM to F and R100-1 TraM binding sites. To test F TraM's ability to bind to R100-1 TraM binding sites EMSA experiments were performed with F and R100-1 binding sites supplied by pNY300 and pRF105, respectively. Approximately 2 μg of pNY300 was digested with *Dra*I and *Sal*I in H Buffer (Roche) in a 40 μL volume for 1 hour at 37$^{\text{O}}$C. Approximately 2 μg of pRF105 was digested with *Dra*I and *Nsi*I in H Buffer (Roche) in a 40 μL volume for 1 hour at 37$^{\text{O}}$C. 2 μL aliquots of each digest were then used during EMSA with or without TraM present. Binding reactions were carried out in TED buffer (50 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 1 mM DTT) in a final

Figure 8.2. Detection of circular *sbmAB* species using non-denaturing gels. Restriction enzymes used to digest PCR products are shown above each lane. Reactions including T4 DNA ligase are designated above each lane. Marker XIV is also shown with the size of each fragment shown to the right. A gel using 850 nM TraM is shown, however, 8.5, 85, and 8 500 nM TraM were also used. Positions of dimers and tetramers of the digested PCR products are shown with arrows.


XhoI
XhoI + ligase
XhoI/DraI
XhoI/DraI + ligase
DraI + ligase
DraI
Marker XIV
1000
900
800
700
600
500
400
300
200
100

volume of 20 µL, and protein was allowed to bind to the DNA for 20 minutes at room temperature. 10 µL of 15 % glycerol was added to each reaction, which were then loaded onto a 5 % polyacrylamide gel and electrophoresed at 30 mA. After running, the gels were stained in ethidium bromide. As shown in Figure 8.3, F TraM does not bind to the R100-1 TraM binding sites (all of which are located in the 495 base pair fragment). TraM does bind to the F TraM binding sites (all of which are found in the 300 base pair fragment). Higher concentrations of TraM were required to see band shifting in comparison to the concentrations required during the determination of association constants. This may be because more DNA was required during these experiments since ethidium bromide, not radioactivity, was used to detect the position of the DNA. Increased DNA concentration can have profound effects on the apparent affinity of protein for DNA and is the reason that radioactivity is more prevalent during the determination of association and dissociation constants. However, the allele specific binding of TraM can still be detected since the F TraM binding sites are fully bound at 4.2 mM, whereas the R100-1 TraM binding sites are not bound until non-specific binding is seen at 210 mM.

Figure 8.3. EMSA of F (pNY300) and R100-1 (pRF105) TraM binding sites. pNY300 was digested with *Dra*I and *Sal*I and pRF105 was digested with *Dra*I and *Nsi*I before use in the experiment. Approximate sizes of the DNA fragments are indicated to the right and left of the gel and the fragment containing the TraM binding sites is designated with an asterisk. Following electrophoresis at 30 mA in a 5 % polyacrylamide gel, the position of DNA fragments was visualized by staining with ethidium bromide.


pNY300
*Dra*I/*Sal*I
pRF105
*Dra*I/*Nsi*I
mM
TraM
0
0.21
0.42
2.1
4.2
21
210
2100
0
2.1
21
105
210
2100
1305
800
690
360
300*
250
1769
1158
690
495*
235

Chapter 9

References

References

Abdel-Monem, M., G. Taucher-Scholz, and M. Q. Klinkert . 1983. Identification of Escherichia coli DNA helicase I as the traI gene product of the F sex factor. Proc. Natl. Acad. Sci. U. S. A **80**:4659-4663.

Abo, T., S. Inamoto, and E. Ohtsubo. 1991. Specific DNA binding of the TraM protein to the *oriT* region of plasmid R100. J. Bacteriol. **173**:6347-6354.

Abo, T. and E. Ohtsubo. 1993. Repression of the *traM* gene of plasmid R100 by its own product and integration host factor at one of the two promoters. J. Bacteriol. **175**:4466-4474.

Abo, T. and E. Ohtsubo. 1995. Characterization of the functional sites in the *oriT* region involved in DNA transfer promoted by sex factor plasmid R100. J. Bacteriol. **177**:4350-4355.

Achtman, M., N. Willetts, and A. J. Clark. 1971. Beginning a genetic analysis of conjugational transfer determined by the F factor in Escherichia coli by isolation and characterization of transfer-deficient mutants. J. Bacteriol. **106**:529-538.

Achtman, M., N. Willetts, and A. J. Clark. 1972. Conjugational complementation analysis of transfer-deficient mutants of F*lac* in Escherichia coli. J. Bacteriol. **110**:831-842.

Achtman, M., N. Kennedy, and R. Skurray. 1977. Cell--cell interactions in conjugating Escherichia coli: role of traT protein in surface exclusion. Proc. Natl. Acad. Sci. U. S. A **74**:5104-5108.

Achtman, M., P. A. Manning, C. Edelbluth, and P. Herrlich. 1979. Export without proteolytic processing of inner and outer membrane proteins encoded by F sex factor *tra* cistrons in Escherichia coli minicells. Proc. Natl. Acad. Sci. U S A **76**:4837-4841.

Alberti, S., S. Oehler, B. Wilcken-Bergmann, and B. Muller-Hill. 1993. Genetic analysis of the leucine heptad repeats of Lac repressor: evidence for a 4-helical bundle. EMBO J. **12**:3227-3236.

Anthony, K. G. 1998. The Role of the F Pilus in Conjugation.
Ref Type: Thesis/Dissertation

Anthony, K. G., W. A. Klimke, J. Manchak, and L. S. Frost. 1999. Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. J. Bacteriol. **181**:5149-5159.

Asada, K., K. Sugimoto, A. Oka, M. Takanami, and Y. Hirota. 1982. Structure of replication origin of the Escherichia coli K-12 chromosome: the presence of spacer sequences in the ori region carrying information for autonomous replication. Nucleic Acids Res. **10**:3745-3754.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and J. A. Struhl (ed). 1987. Current protocols in molecular biology, and supplements. John Wiley and Sons, Inc., New York, N.Y.

Baas, P. D. 1987. Mutational analysis of the bacteriophage phi X174 replication origin. J. Mol. Biol. **198**:51-61.

Baas, P. D. and H. S. Jansz. 1988. Single-stranded DNA phage origins. Curr. Top. Microbiol. Immunol. **136**:31-70.

Baker, T. A. and S. H. Wickner. 1992. Genetics and enzymology of DNA replication in Escherichia coli. Annu. Rev. Genet. **26**:447-477.

Balasubramanian, B., W. K. Pogozelski, and T. D. Tullius. 1998. DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone. Proc. Natl. Acad. Sci. U S A **95**:9738-9743.

Balis, E., A. C. Vatopoulos, M. Kanelopoulou, E. Mainas, G. Hatzoudis, V. Kontogianni, H. Malamou-Lada, S. Kitsou-Kiriakopoulou, and V. Kalapothaki. 1996. Indications of in vivo transfer of an epidemic R plasmid from Salmonella enteritidis to Escherichia coli of the normal human gut flora. J. Clin. Microbiol. **34**:977-979.

Balzer, D., W. Pansegrau, and E. Lanka. 1994. Essential motifs of relaxase (TraI) and TraG proteins involved in conjugative transfer of plasmid RP4. J. Bacteriol. **176**:4285-4295.

Bartel, P., C. T. Chien, R. Sternglanz, and S. Fields. 1993. Elimination of false positives that arise in using the two-hybrid system. Biotechniques **14**:920-924.

Birnboim, H. C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513-1523.

Bramhill, D. and A. Kornberg. 1988. Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome. Cell **52**:743-755.

Brantl, S. 1994. The copR gene product of plasmid pIP501 acts as a transcriptional repressor at the essential repR promoter. Mol. Microbiol. **14**:473-483.

Brantl, S. and E. G. Wagner. 1996. An unusually long-lived antisense RNA in plasmid copy number control: in vivo RNAs encoded by the streptococcal plasmid pIP501. J. Mol. Biol. **255**:275-288.

Brenowitz, M., N. Mandal, A. Pickar, E. Jamison, and S. Adhya. 1991. DNA-binding properties of a lac repressor mutant incapable of forming tetramers. J. Biol. Chem. **266**:1281-1288.

Bruand, C., E. Le Chatelier, S. D. Ehrlich, and L. Janniere. 1993. A fourth class of theta-replicating plasmids: the pAM beta 1 family from gram-positive bacteria. Proc. Natl. Acad. Sci. U. S. A **90**:11668-11672.

Bush, S. M., S. Folta, and D. A. Lannigan. 1996. Use of the yeast one-hybrid system to screen for mutations in the ligand-binding domain of the estrogen receptor. Steroids **61**:102-109.

Byrd, D. R. and S. W. Matson. 1997. Nicking by transesterification: the reaction catalysed by a relaxase. Mol. Microbiol. **25**:1011-1022.

Carey, J. 1988. Gel retardation at low pH resolves trp repressor-DNA complexes for quantitative study. Proc. Natl. Acad. Sci. U. S. A **85**:975-979.

Carey, J. 1991. Gel retardation. Methods Enzymol. **208**:103-117.

Carlson, N. G. and J. W. Little. 1993. Highly cooperative DNA binding by the coliphage HK022 repressor. J. Mol. Biol. **230**:1108-1130.

Cayrol, C., G. Cabrolier, and B. Ducommun. 1997. Use of the two-hybrid system to identify protein-protein interaction temperature-sensitive mutants: application to the CDK2/p21Cip1 interaction. Nucleic Acids Res. **25**:3743-3744.

Chatterjee, J., C. M. Miyamoto, and E. A. Meighen. 1996. Autoregulation of luxR: the Vibrio harveyi lux-operon activator functions as a repressor. Mol. Microbiol. **20**:415-425.

Chevray, P. M. and D. Nathans. 1992. Protein interaction cloning in yeast: identification of mammalian proteins that react with the leucine zipper of Jun. Proc. Natl. Acad. Sci. U. S. A **89**:5789-5793.

Chien, C. T., P. L. Bartel, R. Sternglanz, and S. Fields. 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proc. Natl. Acad. Sci. U. S. A **88**:9578-9582.

Chilley, P. M. and B. M. Wilkins. 1995. Distribution of the ardA family of antirestriction genes on conjugative plasmids. Microbiology **141 (Pt 9)**:2157-2164.

Couturier, M., F. Bex, P. L. Bergquist, and W. K. Maas . 1988. Identification and classification of bacterial plasmids. Microbiol. Rev. **52**:375-395.

Dash, P. K., B. A. Traxler, M. M. Panicker, D. D. Hackney, and E. G. J. Minkley. 1992. Biochemical characterization of Escherichia coli DNA helicase I. Mol. Microbiol. **6**:1163-1172.

Davison, J. 1999. Genetic exchange between bacteria in the environment. Plasmid **42**:73-91.

del Solar, G., M. Moscoso, and M. Espinosa. 1993. In vivo definition of the functional origin of replication (ori(+)) of the promiscuous plasmid pLS1. Mol. Gen. Genet. **237**:65-72.

del Solar, G., M. Moscoso, and M. Espinosa. 1993. Rolling circle-replicating plasmids from gram-positive and gram- negative bacteria: a wall falls. Mol. Microbiol. **8**:789-796.

del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas. 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. **62**:434-464.

del Solar, G. H., A. Puyet, and M. Espinosa. 1987. Initiation signals for the conversion of single stranded to double stranded DNA forms in the streptococcal plasmid pLS1. Nucleic Acids Res. **15**:5561-5580.

Dempsey, W. B. and N. S. Willetts. 1976. Plasmid co-integrates of prophage lambda and R factor R100. J. Bacteriol. **126**:166-176.

Dempsey, W. B. and B. E. Fee. 1990. Integration host factor affects expression of two genes at the conjugal transfer origin of plasmid R100. Mol. Microbiol. **4**:1019-1028.

Dempsey, W. B. 1994. Regulation of R100 conjugation requires traM in cis to traJ. Mol. Microbiol. **13**:987-1000.

Di Laurenzio, L., L. S. Frost, B. B. Finlay, and W. Paranchych. 1991. Characterization of the *oriT* region of the IncFV plasmid pED208. Mol. Microbiol. **5**:1779-1790.

Di Laurenzio, L., L. S. Frost, and W. Paranchych. 1992. The TraM protein of the conjugative plasmid F binds to the origin of transfer of the F and ColE1 plasmids. Mol. Microbiol. **6**:2951-2959.

Di Laurenzio, L. 1992. Protein-DNA interactions at *oriT* Regions.
Ref Type: Thesis/Dissertation

Di Laurenzio, L., D. G. Scraba, W. Paranchych, and L. S. Frost. 1995. Studies on the binding of integration host factor (IHF) and TraM to the origin of transfer of the IncFV plasmid pED208. Mol. Gen. Genet. **247**:726-734.

Disque-Kochem, C. and B. Dreiseikelmann. 1997. The cytoplasmic DNA-binding protein TraM binds to the inner membrane protein TraD in vitro. J. Bacteriol. **179**:6133-6137.

Dixon, W. J., J. J. Hayes, J. R. Levin, M. F. Weidner, B. A. Dombroski, and T. D. Tullius. 1991. Hydroxyl radical footprinting. Methods Enzymol. **208**:380-413.

Eisenberg, S. and A. Kornberg. 1979. Purification and characterization of phiX174 gene A protein. A multifunctional enzyme of duplex DNA replication. J. Biol. Chem. **254**:5328-5332.

Espinosa, M., S. Cohen, M. Couturier, G. del Solar, R. Diaz-Orejas, R. Giraldo, L. Janniere, C. Miller, M. Osborn, and C. M. Thomas. 2000. Plasmid Replication and Copy Number Control, p. 1-47. *In* C. M. Thomas (ed.), The Horizontal Gene Pool: Bacterial plasmids and gene spread. Harwood Academic Publishers, Amsterdam.

Everett, R. and N. Willetts. 1980. Characterisation of an *in vivo* system for nicking at the origin of conjugal DNA transfer of the sex factor F. J. Mol. Biol. **136**:129-150.

Fekete, R. A. and L. S. Frost. 2000. Mobilization of chimeric oriT plasmids by F and R100-1: role of relaxosome formation in defining plasmid specificity. J. Bacteriol. **182**:4022-4027.

Fields, S. and O. Song. 1989. A novel genetic system to detect protein-protein interactions. Nature **340**:245-246.

Finlay, B. B., L. S. Frost, and W. Paranchych. 1986a. Origin of transfer of IncF plasmids and nucleotide sequences of the type II oriT, traM, and traY alleles from ColB4-K98 and the type IV traY allele from R100-1. J. Bacteriol. **168**:132-139.

Finlay, B. B., L. S. Frost, and W. Paranchych. 1986b. Nucleotide sequences of the R1-19 plasmid transfer genes traM, finP, traJ, and traY and the traYZ promoter. J. Bacteriol. **166**:368-374.

Firth, N. and R. Skurray. 1992. Characterization of the F plasmid bifunctional conjugation gene, *traG*. Mol. Gen. Genet. **232**:145-153.

Fowler, T. and R. Thompson. 1986. Shadow promoters in the F plasmid transfer operon. Mol. Gen. Genet. **202**:509-511.

Franke, E. K., H. E. Yuan, K. L. Bossolt, S. P. Goff, and J. Luban. 1994. Specificity and sequence requirements for interactions between various retroviral Gag proteins. J. Virol. **68**:5300-5305.

Freifelder, D. 1982. Physical Biochemistry: Applications to Biochemistry and Molecular Biology. W.H. Freeeman and Company, San Francisco.

Friedman, A. M., T. O. Fischmann, and T. A. Steitz. 1995. Crystal structure of lac repressor core tetramer and its implications for DNA looping. Science **268**:1721-1727.

Frost, L., S. Lee, N. Yanchar, and W. Paranchych. 1989. *finP* and *fisO* mutations in FinP anti-sense RNA suggest a model for FinOP action in the repression of bacterial conjugation by the F*lac* plasmid JCFL0. Mol. Gen. Genet. **218**:152-160.

Frost, L. S., W. Paranchych, and N. S. Willetts. 1984. DNA sequence of the F traALE region that includes the gene for F pilin. J. Bacteriol. **160**:395-401.

Frost, L. S., K. Ippen-Ihler, and R. A. Skurray. 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol. Rev. **58**:162-210.

Frost, L. S., R. A. Fekete, S. S. Penfold, and J. Manchak. 1997. Factors that control TraM expression: The level of TraM defines F plasmid transfer proficiency. Plasmid **37**:220-221.

Frost, L. S. and J. Manchak. 1998. F- phenocopies: characterization of expression of the F transfer region in stationary phase. Microbiology **144 (Pt 9)**:2579-2587.

Fu, Y. H., M. M. Tsai, Y. N. Luo, and R. C. Deonier. 1991. Deletion analysis of the F plasmid *oriT* locus. J. Bacteriol. **173**:1012-1020.

Fukuda, H. and E. Ohtsubo. 1997. Roles of TraI protein with activities of cleaving and rejoining the single-stranded DNA in both initiation and termination of conjugal DNA transfer. Genes Cells **2**:735-751.

Furste, J. P., W. Pansegrau, G. Ziegelin, M. Kroger, and E. Lanka. 1989. Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. Proc. Natl. Acad. Sci. U. S. A **86**:1771-1775.

Gordon, G. S., D. Sitnikov, C. D. Webb, A. Teleman, A. Straight, R. Losick, A. W. Murray, and A. Wright. 1997. Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. Cell **90**:1113-1121.

Grandoso, G., P. Avila, A. Cayon, M. A. Hernando, M. Llosa, and C. F. de La. 2000. Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. J. Mol. Biol. **295**:1163-1172.

Graus-Goldner, A., H. Graus, T. Schlacher, and G. Hogenauer. 1990. The sequences of genes bordering oriT in the enterotoxin plasmid P307: comparison with the sequences of plasmids F and R1. Plasmid **24**:119-131.

Gruss, A. and S. D. Ehrlich. 1989. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. Microbiol. Rev. **53**:231-241.

Guzman, L. M. and M. Espinosa. 1997. The mobilization protein, MobM, of the streptococcal plasmid pMV158 specifically cleaves supercoiled DNA at the plasmid oriT. J. Mol. Biol. **266**:688-702.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557-580.

Hanai, R. and J. C. Wang. 1993. The mechanism of sequence-specific DNA cleavage and strand transfer by phi X174 gene A* protein. J. Biol. Chem. **268**:23830-23836.

Heidekamp, F., P. D. Baas, and H. S. Jansz. 1982. Nucleotide sequences at the phi X gene A protein cleavage site in replicative form I DNAs of bacteriophages U3, G14, and alpha 3. J. Virol. **42**:91-99.

Helinski, D. R., A. E. Toukdarian, and R. P. Novick. 1996. Replication control and other stable maintanance mechanisms of plasmids, p. 2295-2318. *In* F. C. Neidhardt (ed.), Escherichia coli and Salmonella:Cellular and molecular biology. ASM Press, Washington, D.C.

Hertzberg, R. P. and P. B. Dervan. 1984. Cleavage of DNA with methidiumpropyl-EDTA-iron(II): reaction conditions and product analyses. Biochemistry **23**:3934-3945.

Hou, C. L., C. Tang, S. R. Roffler, and T. K. Tang. 2000. Protein 4.1R binding to eIF3-p44 suggests an interaction between the cytoskeletal network and the translation apparatus. Blood **96**:747-753.

Howard, M. T., W. C. Nelson, and S. W. Matson. 1995. Stepwise assembly of a relaxosome at the F plasmid origin of transfer. J. Biol. Chem. **270**:28381-28386.

Howe, I., D. H. Williams, and R. D. Bowen. 1981. Mass Spectrometry: Principles and Applications. McGraw-Hill International Book Company, New York.

Hu, J. C., M. G. Kornacker, and A. Hochschild. 2000. Escherichia coli one- and two-hybrid systems for the analysis and identification of protein-protein interactions. Methods **20**:80-94.

Ikeda, J. E., A. Yudelevich, and J. Hurwitz. 1976. Isolation and characterization of the protein coded by gene A of bacteriophage phiX174 DNA. Proc. Natl. Acad. Sci. U. S. A **73**:2669-2673.

Inamoto, S., Y. Yoshioka, and E. Ohtsubo. 1988. Identification and characterization of the products from the traJ and traY genes of plasmid R100. J. Bacteriol. **170**:2749-2757.

Inamoto, S. and E. Ohtsubo. 1990. Specific binding of the TraY protein to *oriT* and the promoter region for the *traY* gene of plasmid R100. J. Biol. Chem. **265**:6461-6466.

Inamoto, S., T. Abo, and E. Ohtsubo. 1990. Binding sites of integration host factor in *oriT* of plasmid R100. J. Gen. Appl. Microbiol. **36**:287-293.

Inamoto, S., Y. Yoshioka, and E. Ohtsubo. 1991. Site- and strand-specific nicking in vitro at *oriT* by the TraY-TraI endonuclease of plasmid R100. J. Biol. Chem. **266**:10086-10092.

Inamoto, S., H. Fukuda, T. Abo, and E. Ohtsubo. 1994. Site- and strand-specific nicking at oriT of plasmid R100 in a purified system: enhancement of the nicking activity of TraI (helicase I) with TraY and IHF. J. Biochem. (Tokyo) **116**:838-844.

Ingmer, H., E. L. Fong, and S. N. Cohen. 1995. Monomer-dimer equilibrium of the pSC101 RepA protein. J. Mol. Biol. **250**:309-314.

Iordanescu, S. 1993. Characterization of the Staphylococcus aureus chromosomal gene pcrA, identified by mutations affecting plasmid pT181 replication. Mol. Gen. Genet. **241**:185-192.

Iwabuchi, K., B. Li, P. Bartel, and S. Fields. 1993. Use of the two-hybrid system to identify the domain of p53 involved in oligomerization. Oncogene **8**:1693-1696.

James, P., J. Halladay, and E. A. Craig. 1996. Genomic libraries and a host strain designed for highly efficient two- hybrid selection in yeast. Genetics **144**:1425-1436.

Janniere, L., V. Bidnenko, S. McGovern, S. D. Ehrlich, and M. A. Petit. 1997. Replication terminus for DNA polymerase I during initiation of pAM beta 1 replication: role of the plasmid-encoded resolution system. Mol. Microbiol. **23**:525-535.

Ji, T. H. 1983. Bifunctional reagents. Methods Enzymol. **91**:580-609.

Jin, R., X. Zhou, and R. P. Novick. 1996. The inactive pT181 initiator heterodimer, RepC/C, binds but fails to induce melting of the plasmid replication origin. J. Biol. Chem. **271**:31086-31091.

Jin, R., M. E. Fernandez-Beros, and R. P. Novick. 1997. Why is the initiation nick site of an AT-rich rolling circle plasmid at the tip of a GC-rich cruciform? EMBO J. **16**:4456-4466.

Kalpana, G. V. and S. P. Goff. 1993. Genetic analysis of homomeric interactions of human immunodeficiency virus type 1 integrase using the yeast two-hybrid system. Proc. Natl. Acad. Sci. U. S. A **90**:10593-10597.

Kato, G. J., W. M. Lee, L. L. Chen, and C. V. Dang. 1992. Max: functional domains and interaction with c-Myc. Genes Dev. **6**:81-92.

Keegan, L., G. Gill, and M. Ptashne. 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science **231**:699-704.

Kersten, S., D. Kelleher, P. Chambon, H. Gronemeyer, and N. Noy. 1995. Retinoid X receptor alpha forms tetramers in solution. Proc. Natl. Acad. Sci. U. S. A **92**:8645-8649.

Kim, B. and J. W. Little. 1992. Dimerization of a specific DNA-binding protein on the DNA. Science **255**:203-206.

Kim, J., C. Zwieb, C. Wu, and S. Adhya. 1989. Bending of DNA by gene-regulatory proteins: construction and use of a DNA bending vector. Gene **85**:15-23.

Kingsman, A. and N. Willetts. 1978. The requirements for conjugal DNA synthesis in the donor strain during flac transfer. J. Mol. Biol. **122**:287-300.

Klimke, W. A. and L. S. Frost. 1998. Genetic analysis of the role of the transfer gene, traN, of the F and R100-1 plasmids in mating pair stabilization during conjugation. J. Bacteriol. **180**:4036-4043.

Koo, H. S. and D. M. Crothers. 1988. Calibration of DNA curvature and a unified description of sequence- directed bending. Proc. Natl. Acad. Sci. U. S. A **85**:1763-1767.

Koonin, E. V. and T. V. Ilyina. 1993. Computer-assisted dissection of rolling circle DNA replication. Biosystems **30**:241-268.

Kornberg and T. A. Baker. 1992. DNA Replication. W.H. Freeman and Company, New York.

Koronakis, V. and G. Hogenauer. 1986. The sequences of the traJ gene and the 5' end of the traY gene of the resistance plasmid R1. Mol. Gen. Genet. **203**:137-142.

Kramer, H., M. Niemoller, M. Amouyal, B. Revet, B. Wilcken-Bergmann, and B. Muller-Hill. 1987. lac repressor forms loops with linear DNA carrying two suitably spaced lac operators. EMBO J. **6**:1481-1491.

Kruse, H. and H. Sorum. 1994. Transfer of multiple drug resistance plasmids between populations of *Bradyrhizobia* in non-sterile soil. Appl. Environ. Microbiol. **60**:4015-4021.

Kupelwieser, G., M. Schwab, G. Hogenauer, G. Koraimann, and E. L. Zechner. 1998. Transfer Protein TraM Stimulates TraI-Catalyzed Cleavage of the Transfer Origin of Plasmid R1 *in vivo*. J. Mol. Biol. **275**:81-94.

Lahue, E. E. and S. W. Matson. 1990. Purified Escherichia coli F-factor TraY protein binds *oriT*. J. Bacteriol. **172**:1385-1391.

Lane, D., P. Prentki, and M. Chandler. 1992. Use of gel retardation to analyze protein-nucleic acid interactions. Microbiol. Rev. **56**:509-528.

Lanka, E. and B. M. Wilkins. 1995. DNA processing reactions in bacterial conjugation. Annu. Rev. Biochem. **64**:141-169.

Latham, J. A. and T. R. Cech. 1989. Defining the inside and outside of a catalytic RNA molecule. Science **245**:276-282.

Leanna, C. A. and M. Hannink. 1996. The reverse two-hybrid system: a genetic scheme for selection against specific protein/protein interactions. Nucleic Acids Res. **24**:3341-3347.

Lee, M. H., N. Kosuk, J. Bailey, B. Traxler, and C. Manoil. 1999. Analysis of F factor TraD membrane topology by use of gene fusions and trypsin-sensitive insertions. J. Bacteriol. **181**:6108-6113.

Lee, S. H., L. S. Frost, and W. Paranchych. 1992. FinOP repression of the F plasmid involves extension of the half-life of FinP antisense RNA by FinO. Mol. Gen. Genet. **235**:131-139.

Lei, S., L. Pulakat, and N. Gavini. 1999. Genetic analysis of nif regulatory genes by utilizing the yeast two- hybrid system detected formation of a NifL-NifA complex that is implicated in regulated expression of nif genes. J. Bacteriol. **181**:6535-6539.

Lewin, B. 1994. Genes V. Oxford University Press, Oxford.

Lewis, M., G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. A. Schumacher, R. G. Brennan, and P. Lu. 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer [see comments]. Science **271**:1247-1254.

Licitra, E. J. and J. O. Liu. 1996. A three-hybrid system for detecting small ligand-protein receptor interactions. Proc. Natl. Acad. Sci. U. S. A **93**:12817-12821.

Llosa, M., G. Grandoso, and F. de la Cruz. 1995. Nicking activity of TrwC directed against the origin of transfer of the IncW plasmid R388. J. Mol. Biol. **246**:54-62.

Llosa, M., G. Grandoso, M. A. Hernando, and F. de la Cruz. 1996. Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. J. Mol. Biol. **264**:56-67.

Lorenz, M. G. and W. Wackernagel. 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Rev. **58**:563-602.

Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**:265-275.

Luban, J., K. L. Bossolt, E. K. Franke, G. V. Kalpana, and S. P. Goff. 1993. Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. Cell **73**:1067-1078.

Luo, Y., Q. Gao, and R. C. Deonier. 1994. Mutational and physical analysis of F plasmid *traY* protein binding to *oriT*. Mol. Microbiol. **11**:459-469.

Ma, J. and M. Ptashne. 1987a. A new class of yeast transcriptional activators. Cell **51**:113-119.

Ma, J. and M. Ptashne. 1987b. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell **48**:847-853.

Ma, Q. and J. P. Whitlock, Jr. 1997. A novel cytoplasmic protein that interacts with the Ah receptor, contains tetratricopeptide repeat motifs, and augments the transcriptional response to 2,3,7,8-tetrachlorodibenzo-p-dioxin. J. Biol. Chem. **272**:8878-8884.

Maneewannakul, K., P. Kathir, S. Endley, D. Moore, J. Manchak, L. Frost, and K. Ippen-Ihler. 1996. Construction of derivatives of the F plasmid pOX-*tra*715: characterization of *traY* and *traD* mutants that can be complemented *in trans*. Mol. Microbiol. **22**:197-205.

Maneewannakul, S., P. Kathir, and K. Ippen-Ihler. 1992. Characterization of the F plasmid mating aggregation gene traN and of a new F transfer region locus trbE. J. Mol. Biol. **225**:299-311.

Manning, P. A., G. Morelli, and M. Achtman. 1981. traG protein of the F sex factor of Escherichia coli K-12 and its role in conjugation. Proc. Natl. Acad. Sci. U. S. A **78**:7487-7491.

Margulies, C. and J. M. Kaguni. 1996. Ordered and sequential binding of DnaA protein to oriC, the chromosomal origin of Escherichia coli. J. Biol. Chem. **271**:17035-17040.

Marians, K. J. 1996. Replication fork propogation, p. 749-763. *In* F. C. Neidhardt (ed.), Escherichia coli and Salmonella: cellular and molecular biology. ASM Press, Washington, D.C.

Marini, J. C., S. D. Levene, D. M. Crothers, and P. T. Englund. 1982. Bent helical structure in kinetoplast DNA. Proc. Natl. Acad. Sci. U S A **79**:7664-7668.

Masai, H., N. Nomura, Y. Kubota, and K. Arai. 1990. Roles of phi X174 type primo. J. Biol. Chem. **265**:15124-15133.

Masson, L. and D. S. Ray. 1988. Mechanism of autonomous control of the Escherichia coli F plasmid: purification and characterization of the repE gene product. Nucleic Acids Res. **16**:413-424.

Mathews, C. K. and K. E. van Holde. 1990. Information Copying: Replication Biochemistry. Benjamin/Cummings Publishing Company, Redwood City, CA.

Matson, S. W. and B. S. Morton. 1991. Escherichia coli DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. J. Biol. Chem. **266**:16232-16237.

Matson, S. W., W. C. Nelson, and B. S. Morton. 1993. Characterization of the reaction product of the oriT nicking reaction catalyzed by Escherichia coli DNA helicase I. J. Bacteriol. **175** :2599-2606.

Moncalian, G., M. Valle, J. M. Valpuesta, and C. F. de La. 1999. IHF protein inhibits cleavage but not assembly of plasmid R388 relaxosomes. Mol. Microbiol. **31**:1643-1652.

Moore, D., B. A. Sowa, and K. Ippen-Ihler. 1981. Location of an F-pilin pool in the inner membrane. J. Bacteriol. **146**:251-259.

Moore, D., J. H. Wu, P. Kathir, C. M. Hamilton, and K. Ippen-Ihler. 1987. Analysis of transfer genes and gene products within the *traB-traC* region of the Escherichia coli fertility factor, F. J. Bacteriol. **169** :3994-4002.

Mori, M., T. Konno, T. Ozawa, M. Murata, K. Imoto, and K. Nagayama. 2000. Novel interaction of the voltage-dependent sodium channel (VDSC) with calmodulin: does VDSC acquire calmodulin-mediated Ca2+-sensitivity? Biochemistry **39**:1316-1323.

Moscoso, M., G. del Solar, and M. Espinosa. 1995. Specific nicking-closing activity of the initiator of replication protein RepB of plasmid pMV158 on supercoiled or single-stranded DNA. J. Biol. Chem. **270**:3772-3779.

Moscoso, M., R. Eritja, and M. Espinosa. 1997. Initiation of replication of plasmid pMV158: mechanisms of DNA strand- transfer reactions mediated by the initiator RepB protein. J. Mol. Biol. **268**:840-856.

Muniesa, M. and J. Jofre. 1998. Abundance in sewage of bacteriophages that infect Escherichia coli O157:H7 and that carry the Shiga toxin 2 gene. Appl. Environ. Microbiol. **64**:2443-2448.

Nelson, W. C., B. S. Morton, E. E. Lahue, and S. W. Matson. 1993. Characterization of the Escherichia coli F factor traY gene product and its binding sites. J. Bacteriol. **175**:2221-2228.

Nelson, W. C., M. T. Howard, J. A. Sherman, and S. W. Matson. 1995. The *traY* gene product and integration host factor stimulate Escherichia coli DNA helicase I-catalyzed nicking at the F plasmid *oriT*. J. Biol. Chem. **270**:28374-28380.

Nelson, W. C. and S. W. Matson. 1996. The F plasmid traY gene product binds DNA as a monomer or a dimer: structural and functional implications. Mol. Microbiol. **20**:1179-1187.

Nichols, J. C. and K. S. Matthews. 1997. Combinatorial mutations of lac repressor. Stability of monomer-monomer interface is increased by apolar substitution at position 84. J. Biol. Chem. **272**:18550-18557.

Noirot-Gros, M. F., V. Bidnenko, and S. D. Ehrlich. 1994. Active site of the replication protein of the rolling circle plasmid pC194. EMBO J. **13**:4412-4420.

Noirot-Gros, M. F. and S. D. Ehrlich. 1996. Change of a catalytic reaction carried out by a DNA replication protein. Science **274**:777-780.

Noirot, P., J. Bargonetti, and R. P. Novick. 1990. Initiation of rolling-circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin. Proc. Natl. Acad. Sci. U S A **87**:8560-8564.

Noltman, E. A., C. J. Gubler, and S. A. Kuby. 1961. Glucose 6-Phosphate Dehydrogenase (*Zwischenferment*). J. Biol. Chem. **236**:1225-1230.

Novick, R. P. 1989. Staphylococcal plasmids and their replication [published erratum appears in Annu Rev Microbiol 1990;44:10]. Annu. Rev. Microbiol. **43**:537-565.

Oehler, S., E. R. Eismann, H. Kramer, and B. Muller-Hill. 1990. The three operators of the lac operon cooperate in repression. EMBO J. **9**:973-979.

Oka, A., K. Sugimoto, M. Takanami, and Y. Hirota. 1980. Replication origin of the Escherichia coli K-12 chromosome: the size and structure of the minimum DNA segment carrying the information for autonomous replication. Mol. Gen. Genet. **178**:9-20.

Orchard, K. and G. E. May. 1993. An EMSA-based method for determining the molecular weight of a protein-- DNA complex. Nucleic Acids Res. **21**:3335-3336.

Pal, S. K. and D. K. Chattoraj. 1988. P1 plasmid replication: initiator sequestration is inadequate to explain control by initiator-binding sites. J. Bacteriol. **170**:3554-3560.

Panicker, M. M. and E. G. J. Minkley. 1992. Purification and properties of the F sex factor TraD protein, an inner membrane conjugal transfer protein. J. Biol. Chem. **267**:12761-12766.

Pansegrau, W., D. Balzer, V. Kruft, R. Lurz, and E. Lanka. 1990. In vitro assembly of relaxosomes at the transfer origin of plasmid RP4. Proc. Natl. Acad. Sci. U. S. A **87**:6555-6559.

Pansegrau, W., E. Lanka, P. T. Barth, D. H. Figurski, D. G. Guiney, D. Haas, D. R. Helinski, H. Schwab, V. A. Stanisich, and C. M. Thomas. 1994. Complete nucleotide sequence of Birmingham IncP alpha plasmids. Compilation and comparative analysis. J. Mol. Biol. **239** :623-663.

Pansegrau, W. and E. Lanka. 1996a. Enzymology of DNA transfer by conjugative mechanisms. Prog. Nucleic Acid Res. Mol. Biol. **54**:197-251.

Pansegrau, W. and E. Lanka. 1996b. Mechanisms of initiation and termination reactions in conjugative DNA processing. Independence of tight substrate binding and catalytic activity of relaxase (TraI) of IncPalpha plasmid RP4. J. Biol. Chem. **271**:13068-13076.

Park, K., S. Mukhopadhyay, and D. K. Chattoraj. 1998. Requirements for and regulation of origin opening of plasmid P1. J. Biol. Chem. **273**:24906-24911.

Penfold, S. S., K. Usher, and L. S. Frost. 1994. The nature of the traK4 mutation in the F sex factor of Escherichia coli. J. Bacteriol. **176**:1924-1931.

Penfold, S. S. 1995. Transcriptional Regulation of the F Plasmid.
Ref Type: Thesis/Dissertation

Penfold, S. S., J. Simon, and L. S. Frost. 1996. Regulation of the expression of the *traM* gene of the F sex factor of Escherichia coli. Mol. Microbiol. **20**:549-558.

Perwez, T. and R. Meyer. 1996. MobB protein stimulates nicking at the R1162 origin of transfer by increasing the proportion of complexed plasmid DNA. J. Bacteriol. **178**:5762-5767.

Peters, K. and F. M. Richards. 1977. Chemical cross-linking: reagents and problems in studies of membrane structure. Annu. Rev. Biochem. **46**:523-551.

Phillips, C. A., J. Gordon, and E. K. Spicer. 1996. Bacteriophage T4 regA protein binds RNA as a monomer, overcoming dimer interactions. Nucleic Acids Res. **24**:4319-4326.

Phizicky, E. M. and S. Fields. 1995. Protein-protein interactions: methods for detection and analysis. Microbiol. Rev. **59**:94-123.

Pitkanen, J., V. Doucas, T. Sternsdorf, T. Nakajima, S. Aratani, K. Jensen, H. Will, P. Vahamurto, J. Ollila, M. Vihinen, H. S. Scott, S. E. Antonarakis, J. Kudoh, N. Shimizu, K. Krohn, and P. Peterson. 2000. The autoimmune regulator protein has transcriptional transactivating properties and interacts with the common coactivator CREB-binding protein. J. Biol. Chem. **275**:16802-16809.

Polzleitner, E., E. L. Zechner, W. Renner, R. Fratte, B. Jauk, G. Hogenauer, and G. Koraimann. 1997. TraM of plasmid R1 controls transfer gene expression as an integrated control element in a complex regulatory network. Mol. Microbiol. **25**:495-507.

Pradel, E., C. T. Parker, and C. A. Schnaitman. 1992. Structures of the *rfaB*, *rfaI*, *rfaJ*, and *rfaS* genes of Escherichia coli K-12 and their roles in assembly of the lipopolysaccharide core. J. Bacteriol. **174**:4736-4745.

Price, M. A. and T. D. Tullius. 1992. Using hydroxyl radical to probe DNA structure. Methods Enzymol. **212**:194-219.

Ratnakar, P. V., B. K. Mohanty, M. Lobert, and D. Bastia. 1996. The replication initiator protein pi of the plasmid R6K specifically interacts with the host-encoded helicase DnaB. Proc. Natl. Acad. Sci. U. S. A **93**:5522-5526.

Reeves, P. and N. Willetts. 1974. Plasmid specificity of the origin of transfer of sex factor F. J. Bacteriol. **120**:125-130.

Roth, M. J., D. R. Brown, and J. Hurwitz. 1984. Analysis of bacteriophage phi X174 gene A protein-mediated termination and reinitiation of phi X DNA synthesis. II. Structural characterization of the covalent phi X A protein-DNA complex. J. Biol. Chem. **259**:10556-10568.

Rousseau, P., M. Betermier, M. Chandler, and R. Alazard. 1996. Interactions between the repressor and the early operator region of bacteriophage Mu. J. Biol. Chem. **271**:9739-9745.

Salyers, A. A. and N. B. Shoemaker. 1996. Resistance gene transfer in anaerobes: new insights, new problems. Clin. Infect. Dis. **23 Suppl 1**:S36-S43.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

SantaLucia, J., Jr. 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest- neighbor thermodynamics. Proc. Natl. Acad. Sci. U S A **95**:1460-1465.

Sastre, J. I., E. Cabezon, and F. de la Cruz. 1998. The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. J. Bacteriol. **180**:6039-6042.

Schandel, K. A., M. M. Muller, and R. E. Webster. 1992. Localization of TraC, a protein involved in assembly of the F conjugative pilus. J. Bacteriol. **174**:3800-3806.

Scherzinger, E., V. Haring, R. Lurz, and S. Otto. 1991. Plasmid RSF1010 DNA replication in vitro promoted by purified RSF1010 RepA, RepB and RepC proteins. Nucleic Acids Res. **19**:1203-1211.

Scherzinger, E., R. Lurz, S. Otto, and B. Dobrinski. 1992. In vitro cleavage of do. Nucleic Acids Res. **20**:41-48.

Scherzinger, E., V. Kruft, and S. Otto. 1993. Purification of the large mobilization protein of plasmid RSF1010 and characterization of its site-specific DNA-cleaving/DNA-joining activity. Eur. J. Biochem. **217**:929-938.

Schildbach, J. F., C. R. Robinson, and R. T. Sauer. 1998. Biophysical characterization of the TraY protein of Escherichia coli F factor. J. Biol. Chem. **273**:1329-1333.

Schwab, M., H. Gruber, and G. Hogenauer. 1991. The TraM protein of plasmid R1 is a DNA-binding protein. Mol. Microbiol. **5**:439-446.

Schwab, M., H. Reisenzein, and G. Hogenauer. 1993. TraM of plasmid R1 regulates its own expression. Mol. Microbiol. **7**:795-803.

Sekimizu, K., B. Y. Yung, and A. Kornberg. 1988. The dnaA protein of Escherichia coli. Abundance, improved purification, and membrane binding. J. Biol. Chem. **263**:7136-7140.

Sherman, J. A. and S. W. Matson. 1994. Escherichia coli DNA helicase I catalyzes a sequence-specific cleavage/ligation reaction at the F plasmid origin of transfer. J. Biol. Chem. **269**:26220-26226.

Sikorski, R. S. and J. D. Boeke. 1991. In vitro mutagenesis and plasmid shuffling: from cloned gene to mutant yeast. Methods Enzymol. **194**:302-318.

Silver, P. A., L. P. Keegan, and M. Ptashine. 1984. Amino terminus of the yeast GAL4 gene product is sufficient for nuclear localization. Proc. Natl. Acad. Sci. U. S. A **81**:5951-5955.

Spassky, A. and D. S. Sigman. 1985. Nuclease activity of 1,10-phenanthroline-copper ion. Conformational analysis and footprinting of the lac operon. Biochemistry **24**:8050-8056.

Stockwell, D. T. and W. B. Dempsey. 1997. The finM promoter and the traM promoter are the principal promoters of the traM gene of the antibiotic resistance plasmid R100. Mol. Microbiol. **26**:455-467.

Syvanen, M. 1994. Horizontal gene transfer: evidence and possible consequences. Annu. Rev. Genet. **28**:237-261.

Szpirer, C. Y., M. Faelen, and M. Couturier. 2000. Interaction between the RP4 coupling protein TraG and the pBHR1 mobilization protein Mob [In Process Citation]. Mol. Microbiol. **37**:1283-1292.

Tabor, S. and C. C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. U S A **82**:1074-1078.

Takacs, A. M., T. Das, and A. K. Banerjee. 1993. Mapping of interacting domains between the nucleocapsid protein and the phosphoprotein of vesicular stomatitis virus by using a two-hybrid system. Proc. Natl. Acad. Sci. U. S. A **90**:10375-10379.

Takechi, S., H. Matsui, and T. Itoh. 1995. Primer RNA synthesis by plasmid-specified Rep protein for initiation of ColE2 DNA replication. EMBO J. **14**:5141-5147.

Taki, K., T. Abo, and E. Ohtsubo. 1998. Regulatory mechanisms in expression of the traY-I operon of sex factor plasmid R100: involvement of traJ and traY gene products. Genes Cells **3**:331-345.

Taylor, J. D., I. G. Badcoe, A. R. Clarke, and S. E. Halford. 1991. EcoRV restriction endonuclease binds all DNA sequences with equal affinity. Biochemistry **30**:8743-8753.

te Riele H., B. Michel, and S. D. Ehrlich. 1986. Are single-stranded circles intermediates in plasmid DNA replication? EMBO J. **5**:631-637.

Thomas, C. D., T. T. Nikiforov, B. A. Connolly, and W. V. Shaw. 1995. Determination of sequence specificity between a plasmid replication initiator protein and the origin of replication. J. Mol. Biol. **254**:381-391.

Thompson, J. F. and A. Landy. 1988. Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes. Nucleic Acids Res. **16**:9687-9705.

Thompson, R. and L. Taylor. 1982. Promoter mapping and DNA sequencing of the F plasmid transfer genes traM and traJ. Mol. Gen. Genet. **188**:513-518.

Thompson, T. L., M. B. Centola, and R. C. Deonier. 1989. Location of the nick at *oriT* of the F plasmid. J. Mol. Biol. **207**:505-512.

Top, E., M. Mergeay, D. Springael, and W. Verstraete. 1990. Gene escape model: Transfer of heavy metal resistance genes from *Escherichia coli* to *Alcaligenes eutrophus* on agar plates and in soil samples. Appl. Environ. Microbiol. **56**:2471-2479.

Traxler, B. A. and E. G. Minkley, Jr. 1988. Evidence that DNA helicase I and oriT site-specific nicking are both functions of the F TraI protein. J. Mol. Biol. **204**:205-209.

Tsai, M. M., Y. H. Fu, and R. C. Deonier. 1990. Intrinsic bends and integration host factor binding at F plasmid *oriT*. J. Bacteriol. **172**:4603-4609.

Tullius, T. D. and B. A. Dombroski. 1985. Iron(II) EDTA used to measure the helical twist along any DNA molecule. Science **230**:679-681.

Tullius, T. D. and B. A. Dombroski. 1986. Hydroxyl radical "footprinting": high-resolution information about DNA- protein contacts and application to lambda repressor and Cro protein. Proc. Natl. Acad. Sci. U S A **83**:5469-5473.

Tullius, T. D., B. A. Dombroski, M. E. Churchill, and L. Kam. 1987. Hydroxyl radical footprinting: a high-resolution method for mapping protein-DNA contacts. Methods Enzymol. **155**:537-558.

Uga, H., F. Matsunaga, and C. Wada. 1999. Regulation of DNA replication by iterons: an interaction between the ori2 and incC regions mediated by RepE-bound iterons inhibits DNA replication of mini-F plasmid in Escherichia coli. EMBO J. **18**:3856-3867.

Uga, H., H. Komori, K. Miki, and C. Wada. 2000. The iteron regions necessary for the RepE-iteron interaction in vivo in mini-F plasmids of Escherichia coli. J. Biochem. (Tokyo) **127**:537-541.

Urh, M., D. York, and M. Filutowicz. 1995. Buffer composition mediates a switch between cooperative and independent binding of an initiator protein to DNA. Gene **164**:1-7.

van Biesen, T. and L. S. Frost. 1994. The FinO protein of IncF plasmids binds FinP antisense RNA and its target, traJ mRNA, and promotes duplex formation. Mol. Microbiol. **14**:427-436.

van Mansfeld, A. D., S. A. Langeveld, P. D. Baas, H. S. Jansz, G. A. van der Marel, G. H. Veeneman, and J. H. van Boom. 1980. Recognition sequence of bacteriophage phi X174 gene A protein--an initiator of DNA replication. Nature **288**:561-566.

van Mansfeld, A. D., H. A. van Teeffelen, P. D. Baas, and H. S. Jansz. 1986. Two juxtaposed tyrosyl-OH groups participate in phi X174 gene A protein catalysed cleavage and ligation of DNA. Nucleic Acids Res. **14**:4229-4238.

van, d. M., Jr., W. M. de Vos, S. Harayama, and A. J. Zehnder. 1992. Molecular mechanisms of genetic adaptation to xenobiotic compounds. Microbiol. Rev. **56**:677-694.

Verdino, P., W. Keller, H. Strohmaier, K. Bischof, H. Lindner, and G. Koraimann. 1999. The essential transfer protein TraM binds to DNA as a tetramer. J. Biol. Chem. **274**:37421-37428.

Wang, P. Z., S. J. Projan, V. Henriquez, and R. P. Novick. 1992. Specificity of origin recognition by replication initiator protein in plasmids of the pT181 family is determined by a six amino acid residue element. J. Mol. Biol. **223**:145-158.

Wang, P. Z., S. J. Projan, V. Henriquez, and R. P. Novick. 1993. Origin recognition specificity in pT181 plasmids is determined by a functionally asymmetric palindromic DNA element. EMBO J. **12**:45-52.

Wang, Z. F., M. L. Whitfield, T. C. Ingledue, III, Z. Dominski, and W. F. Marzluff. 1996. The protein that binds the 3' end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. Genes Dev. **10**:3028-3040.

Waters, V. L. and D. G. Guiney. 1993. Processes at the nick region link conjugation, T-DNA transfer and rolling circle replication. Mol. Microbiol. **9**:1123-1130.

Weigel, C., A. Schmidt, B. Ruckert, R. Lurz, and W. Messer. 1997. DnaA protein binding to individual DnaA boxes in the Escherichia coli replication origin, oriC. EMBO J. **16**:6574-6583.

Willetts, N. and M. Achtman. 1972. Genetic analysis of transfer by the *Escherichia coli* sex factor F, using P1 transductional complementation. J. Bacteriol. **110**:843-851.

Willetts, N. 1977. The transcriptional control of fertility in F-like plasmids. J. Mol. Biol. **112**:141-148.

Willetts, N. and B. Wilkins. 1984. Processing of plasmid DNA during bacterial conjugation. Microbiol. Rev. **48**:24-41.

Willetts, N. and J. Maule. 1986. Specificities of IncF plasmid conjugation genes. Genet. Res. **47**:1-11.

Willetts, N. S. and D. J. Finnegan. 1970. Characteristics of *E. coli* K12 strains carrying both an F prime and an R factor. Genet. Res. **16**:113-122.

Willetts, N. S. 1972. Location of the origin of transfer of the sex factor F. J. Bacteriol. **112**:773-778.

Yoshioka, Y., H. Ohtsubo, and E. Ohtsubo. 1987. Repressor gene finO in plasmids R100 and F: constitutive transfer of plasmid F is caused by insertion of IS3 into F finO. J. Bacteriol. **169**:619-623.

Zechner, E. L., F. de la Cruz, R. Eisenbrandt, A. M. Grahn, G. Koraimann, E. Lanka, G. Muth, W. Pansegrau, C. M. Thomas, B. M. Wilkins, and M. Zatyka. 2000. Conjugative-DNA Transfer Processes, p. 87-174. *In* C. M. Thomas (ed.), The Horizontal Gene Pool: Bacterial plasmids and gene spread. Harwood Academic Publishers, Amsterdam.

Zhang, S. and R. Meyer. 1997. The relaxosome protein MobC promotes conjugal plasmid mobilization by extending DNA strand separation to the nick site at the origin of transfer. Mol. Microbiol. **25**:509-516.

Zhao, A. C. and S. A. Khan. 1996. An 18-base-pair sequence is sufficient for termination of rolling- circle replication of plasmid pT181. J. Bacteriol. **178**:5222-5228.

Ziegelin, G., J. P. Furste, and E. Lanka. 1989. TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. J. Biol. Chem. **264**:11989-11994.

Ziegelin, G., W. Pansegrau, R. Lurz, and E. Lanka. 1992. TraK protein of conjugative plasmid RP4 forms a specialized nucleoprotein complex with the transfer origin. J. Biol. Chem. **267**:17279-17286.

Zinkel, S. S. and D. M. Crothers. 1990. Comparative gel electrophoresis measurement of the DNA bend angle induced by the catabolite activator protein. Biopolymers **29**:29-38.